

EGYETEMI DOKTORI (PhD) ÉRTEKEZÉS

Dr. Török Zsolt

**Könnyproteomikai és képfeldolgozási technikák alkalmazhatóságának vizsgálata a
retinopathia diabetica szűrésében.**

DEBRECENI EGYETEM

KLINIKAI ORVOSTUDOMÁNYOK DOKTORI ISKOLA

Debrecen, 2014.

EGYETEMI DOKTORI (Ph.D.) ÉRTEKEZÉS

**Könnyproteomikai és képfeldolgozási technikák alkalmazhatóságának vizsgálata a
retinopathia diabetica szűrésében.**

Dr. Török Zsolt

Témavezető: Dr. Csutak Adrienne



DEBRECENI EGYETEM

Klinikai Orvostudományok Doktori Iskola

Debrecen, 2014

Tartalom

Rövidítések és mozaikszavak jegyzéke	5
1. Bevezetés	7
A témaválasztás indoklása	7
2. Irodalmi áttekintés	8
2.1. A retinopathia diabetica kialakulása	8
2.1.1. A diabetes mellitus	8
2.1.2. A diabetes mellitus krónikus szövődményei	9
2.1.3. Retinopathia diabetica – a diabetes mellitus egyik leggyakoribb szövődménye	10
2.1.4. A retinopathia diabetica okozta látásromlás	12
2.2. A retinopathia diabetica szűrése	14
2.2.1. A retinafényképezés	15
2.2.2. A jelenlegi klinikai gyakorlatban alkalmazott szűrőeljárások bemutatása	17
2.3. Az automatizált szűrést célzó eljárások bemutatása	19
2.3.1. A képfeldolgozás alapú módszerek általában	19
2.3.2. Microaneurysma detektor – a leggyakrabban alkalmazott képfeldolgozás alapú szűrőeljárás	26
2.3.3. Könnyproteomikai alapú eljárások	27
2.3.4. Kombinált eljárás – a könnyproteomika és a képfeldolgozás kombinálása	30
2.4. Célkitűzések	31
3. Metodikák	32
3.1. A könnyproteomikai alapú szűrés módszertanának vizsgálata során alkalmazott módszerek	32
3.1.1. Betegvizsgálat és mintavétel	32
3.1.2. Könnyproteomikai analízis	33
3.1.3. A gépi tanuló eljárások alkalmazása az automatizált könnyproteomikai alapú szűrőrendszerben	34
3.1.4. Az adatelemzés során alkalmazott szoftver eszközök	35
3.2. A kombinált eljárást alkalmazó szűrés módszertanának vizsgálata során alkalmazott módszerek	36
3.2.1. Betegvizsgálat és mintavétel	36
3.2.2. Az alkalmazott képfeldolgozási módszertan bemutatása	37
3.2.3. Könnyproteomikai vizsgálatok	40
3.2.4. A gépi tanuló eljárás alkalmazása	40
3.2.5. Az adatelemzés céljára szolgáló szoftver eszközök	43

3.3. Etikai engedély, betegbeleegyezés	43
3.4. Archiválás	43
4. Eredmények	43
4.1 A könnyproteomikai alapú eljárás értékelésével kapott eredmények.	43
4.1.1. Hat különböző gépi tanuló eljárás összehasonlítása	43
4.1.2. Az adathalmaz méretének hatása a modell teljesítményére	47
4.2. A kombinált módszertan alkalmazásával kapott eredmények	48
5. Megbeszélés.....	50
6. Összefoglalás.....	53
6.1. Magyar nyelvű összefoglaló.....	53
6.2. Angol nyelvű összefoglaló	54
7. Irodalomjegyzék.....	55
7.1. Az értekezésben hivatkozott közlemények listája	55
7.2. Az értekezés alapjául szolgáló saját közlemények listája.....	61
8. Tárgyszavak	61
8.1. Magyar nyelvű tárgyszavak.....	62
8.2. Angol nyelvű tárgyszavak.....	62
9. Köszönetnyilvánítás	62
10. Függelék	63

Rövidítések és mozaikszavak jegyzéke

DM: diabetes mellitus

RD: retinopathia diabetica

MA: microaneurysma

BH: blot hemorrhage (foltszerű bevérvzés)

EX: exudate (exudatum)

IDDM: insulin-dependent diabetes mellitus (insulindependens diabetes mellitus)

NIDDM: non-insulin-dependent diabetes mellitus (nem insulindependens diabetes mellitus)

AMD: age-related macular degeneration (időskori macula degeneratio)

BRB: blood-retinal barrier (vér-retina gát)

LC-MS: liquid chromatography–mass spectrometry (folyadék kromatográfia-tömegspektrometria)

HPLC: high pressureLiquid chromatography (nagy teljesítményű folyadék kromatográfia)

D: dimenzió

kNN: k-Nearest Neighbors (K-legközelebbi szomszéd (algoritmus – gépi tanulás))

SVM: support vector machine (tartóvektor-gép (algoritmus – gépi tanulás))

rpart: recursive partitioning (rekurzív partícionáló (algoritmus – gépi tanulás))

randomForest: Random Forest (”véletlen erdők” (algoritmus – gépi tanulás))

naiveBayes: naive Bayes (naive Bayes (algoritmus – gépi tanulás))

logReg: logistic regression (algoritmus – gépi tanulás)

CLAHE: contrast limited adaptive histogram equalization (kontraszt limitált adaptív hisztogram kiegyenlítés)

GBM: gradient boosting machine (grádiens fokozó gép (gépi tanulás))

RGB: red, green, and blue (vörös, zöld és kék (additív színkeverés modellje))

PCA: principal components analysis (főkomponens analízis)

SENS: sensitivity (szenzitivitás)

SPC: specificity (specifitás)

ACC: accuracy (pontosság)

PREC: positive predictive value (pozitív prediktív érték)

NPV: negative predictive value (negatív prediktív érték)

F1: F-measure (F-érték)

LRP: likelihood ratio positive (pozitív valószínűségi hányados)

LRN: likelihood ratio negative (negatív valószínűségi hányados)

1. Bevezetés

A témaválasztás indoklása

Az értekezés tárgya a diabetes mellitus (DM) szövődményeként fellépő retinopathia diabetica (RD) szűrésére alkalmazható módszertanok vizsgálata. A téma fontosságát az alapbetegség, illetve szövődményeinek magas prevalenciája adja, valamint az a tény, hogy a RD-t, az aktív életkorban bekövetkező látásvesztés leggyakoribb kiváltó okaként tartják nyilván [1].

A RD kezdetben tünetmentes, így a betegek sokszor csak igen későn a betegség előrehaladott stádiumában fordulnak orvoshoz. A RD progressziója korai diagnózissal és kezeléssel jelentősen lassítható, ezért fontos a veszélyeztetett betegcsoport – a DM-ban szenvedő betegek – rendszeres szemészeti szűrővizsgálata.

A RD szűrést napjainkban - ahogy azt alább részletesen bemutatjuk - egészségügyi képesítéssel bíró szakemberek végzik, akik a legtöbb esetben személyesen találkoznak a betegekkel a vizsgálat során. A központosított szűrővizsgálatok esetén a vizsgáló csak a beteg fundusáról készített digitális fotót kapja meg és használja az értékeléshez. Mindkét protokoll hátránya, hogy alkalmazásához jelentős számú, speciálisan képzett szakemberre van szükség, aminek költségvonzata igen magas.

Az utóbbi években számos kezdeményezés indult útjára azzal a céllal, hogy a RD szűrése során hagyományosan alkalmazott protokollokba olyan módszertani lépéseket illesszenek, melyek – támaszkodva a modern informatikai és laboratóriumi technológiákra – csökkentik a szűrővizsgálatok humánerőforrás igényét. A humánerőforrás igény csökkentése a szűrőprogramok költségének csökkenésével járhat, ami a vizsgálatokat a DM-ban szenvedő betegek szélesebb köre számára teheti elérhetővé. Munkacsoportunk ezen új módszertanok közül az alábbiakat választotta vizsgálati célpontjául:

- a digitalizált szemfenéki képeken azonosítható, RD-ra jellemző elváltozások automatizált felismerése a képfeldolgozás eszközeivel
- a könnyben található fehérjék RD hatására bekövetkező, jellemző koncentráció változásának vizsgálata proteomikai eljárással
- a két fenti módszertan kombinálására épülő eljárások.

2. Irodalmi áttekintés

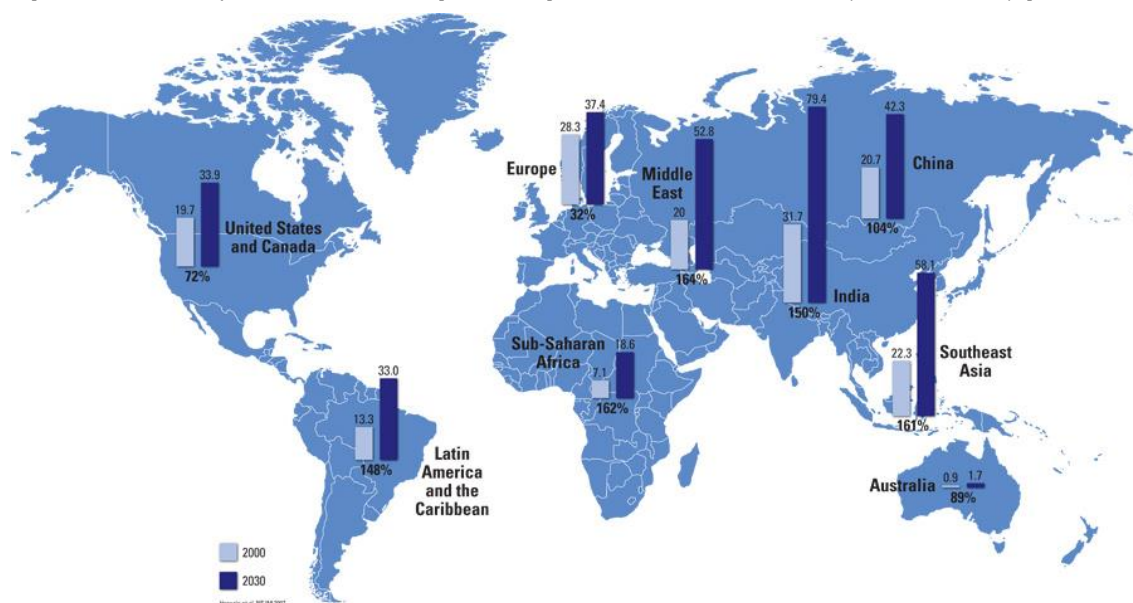
2.1. A retinopathia diabetica kialakulása

2.1.1. A diabetes mellitus

A DM meghatározás az anyagcsere betegségek egy csoportját jelenti, melyek huzamosabban fennálló magas vércukorszinttel jellemezhetőek. Globális prevalenciáját a 20–79 éves korcsoportban 6,4%-osra becsülték 2010-ben, ami 285 millió beteget jelentett. Egyes prognózisok szerint 2030-ra a betegek aránya 7,7%-ra, számuk 439 millióra nőhet. A várható növekedés 69%-os a fejlődő és 20%-os a fejlett országokban, ahogy azt az 1. ábra is szemlélteti [2,3].

Boston Scientific

http://www.bostonscientific-international.com/templatedata/imports/HTML/intl/WorldDiabetesDay/worlddiabetesday_prevalence_intl.html



1. ábra. A cukorbetegség globális prevalenciája 2000-ben, valamint annak várható növekedése 2030-ig.

A cukorbetegségnek két fő típusa különíthető el, melyek közül az I. típus esetén abszolút inzulinhiány alakul ki ismeretlen okból, vagy valamely immunológiai mechanizmus következtében, míg a II. típus esetén a relatív inzulinhiány áll a betegség kialakulásának hátterében. A cukorbetegség nagy hányada II. típusú cukorbetegségben szenved (kb. 90%), kisebb részüket érinti csak az I. típusú cukorbetegség [4].

A DM krónikus betegség, melynek menedzsmentje számos az életmódhoz kapcsolódó előírások mellett többnyire gyógyszeres terápiát is tartalmaz. A kezelések célja a betegek vércukor értékeinek

normál tartományban tartása. A sikeres betegségmenedzsmentnek alapvető szerepe van a DM krónikus mellékhatásainak megelőzésében, súlyosságának csökkentésében. A tapasztalatok szerint a jó terápiás együttműködést mutató betegek esélye az egyes szövődmények kialakulására jelentősen elmarad a nem kellően együttműködő betegekéétől [5,6].

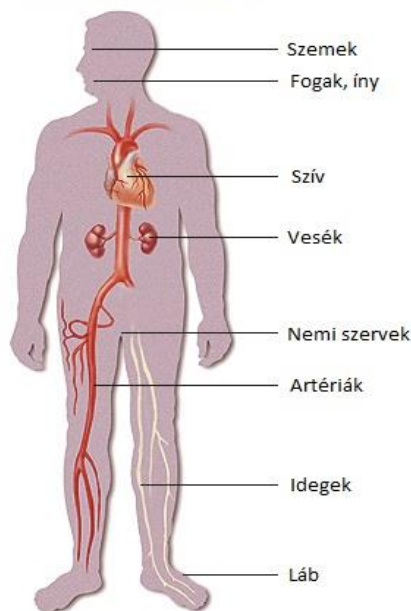
2.1.2. A diabetes mellitus krónikus szövődményei

A DM I-es és II-es típusára egyaránt jellemző, hogy a betegség huzamosabb fennállását követően krónikus szövődmények alakulhatnak ki. Különösen igaz ez azokra az esetekre, amelyeknél a késői diagnózis miatt a betegség sokáig kezeletlen marad. A hosszú távú szövődmények elsősorban a kis és nagyerek érintettségének következményei.

A nagyerek károsodása (macroangiopathia) leggyakrabban az agyi erek érintettségének következtében kialakuló stroke, a coronaria erek károsodása okozta angina, szívinfarktus, valamint a perifériás artériák szűkülete okozta alsó végtagi fájdalom (claudicatio intermittens) és szövetelhalás képében jelentkezik. A kiserek károsodása (microangiopathia), a szervezet egészét érintő megbetegedésről lévén szó, számos szövődmény képében jelentkezhet. Ilyenek például a nephropathia, neuropathia, encephalopathia diabetica, valamint jelen értekezés tárgyához kapcsolódóan a RD (2. ábra).

Eredeti ábra: St Michael's Hospital, Toronto <http://healthtips.stmichaelshospital.com/tag/diabetes-complications/>

A diabetes szövődményei



2. ábra. A DM szövődményeinek leggyakoribb szervi lokalizációja.

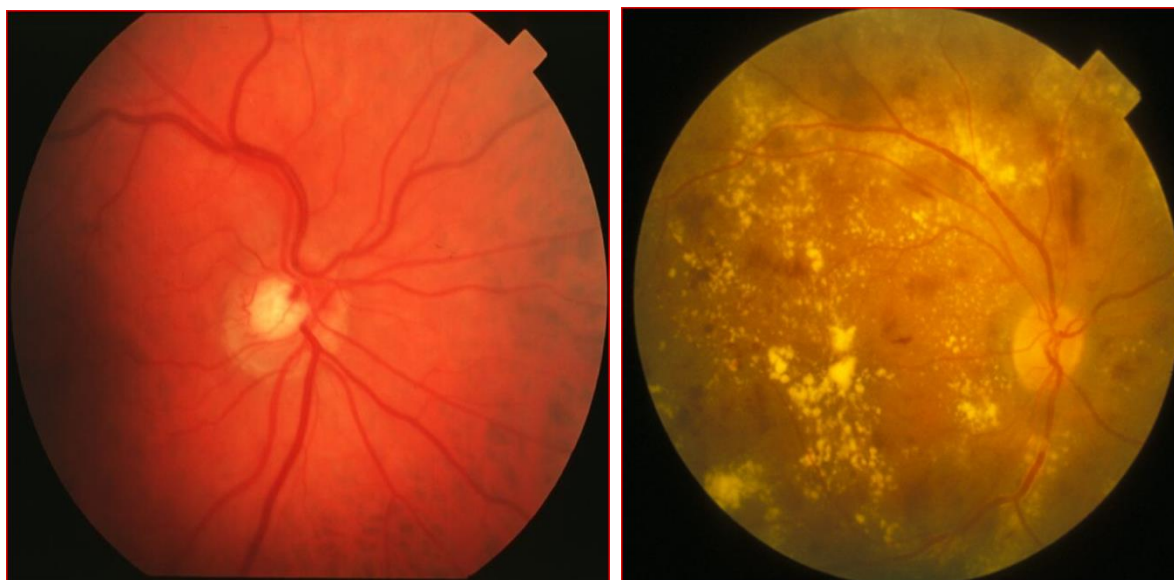
2.1.3. Retinopathia diabetica – a diabetes mellitus egyik leggyakoribb szövődménye

A RD a DM leggyakoribb szövődményei közé tartozik [7]. Kialakulásának hátterében a retina kisereinek károsodása áll. Szövettanilag a membrana basalis megvastagodása a megbetegedés kórjelzője. A megvastagodás mellett a membrán szerkezete is megváltozik, aminek következtében átjárhatóbbá válik bizonyos plazma fehérjék és egyéb szerves makromolekulák számára. A kialakuló eltérések nagymértékben rontják a retina microcirculatioját, emellett a DM okozta anyagcsere eltérések direkt módon is hatva nagymértékű pericyta vesztést okoznak. Ennek következtében a retina erek keringésének autoregulatioja is nagymértékben károsodik.

A RD stádiumai

A RD progressioja során, a különböző szemfenéki elváltozások megjelenése alapján négy stádiumot különböztetnek meg: i.) Praeretinopathia; ii.) Nonproliferatív (háttér) retinopathia; iii.) Preproliferatív retinopathia; iv.) Proliferatív retinopathia. Ilyen elváltozások például a - képfeldolgozási eszközökkel történő detektálás szempontjából legfontosabb - az 1. stádiumban megjelenő, de szemtükörrel (apró vörös pontként) csak a 2. stádiumban láthatóvá váló microaneurysmák (MA-k). A 2. stádiumban emellett intraretinális bevérvések, kemény exudátumok és a macula oedemája is megjelenhet (3. ábra). A pont/folt alakú bevérvések a capillarisk rupturája következtében alakulnak ki és többnyire a MA-nál mélyebb rétegben helyezkednek el. A kemény exudátumokat a megnövekedett membrán permeabilitás következtében kiszivárgó lipidek lerakódása hozza létre. A MA-k idővel thrombotizálnak és együttesen a vér-retina gát (BRB) károsodása okozta ödémával hozzájárulnak a maculopathia diabetica kialakulásához. A RD későbbi stádiumaiban a látható eltérések súlyosbodnak. Jelentős bevérvések, puha exudátumok (gyapottépésszerű gócok) képében jelentkező infarctusos területek, a proliferatív stádiumban pedig neovascularisatio jelei láthatóak [8].

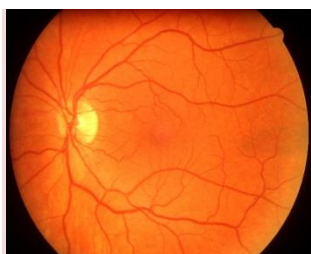
A szemfenéki tünetek felsorolására a teljesség igénye nélkül került sor, azonban fontos hangsúlyozni, hogy jelen munka szempontjából kiemelt szerepe van a MA-nak, mivel ez a laesio a RD legkorábbi jelei közé tartozik. Ha a fundus vizsgálata során MA-t nem tudunk azonosítani a RD diagnózisa megkérdőjelezhető [9,10,11]. Mivel célunk egy szűrés céljára alkalmazható eljárás kidolgozása, a RD-t a lehető legkorábbi stádiumban szeretnénk felismerni és nem törekszünk a betegség egyes fázisainak elkülönítésére. A vizsgálataink során alkalmazott képfeldolgozás alapú eljárás kidolgozásakor ezért helyeztük a hangsúlyt a MA detektálására. Más képfeldolgozó eljárások egyéb laesiók felismerésére, sok esetben több laesio párhuzamos felismerésére is képesek.



3. ábra. Egészséges szemfenék képe balra, RD képe maculopathiával jobbra.

Mivel más megbetegedések is károsíthatják a retinát, el kell különítenünk a RD illetve más retinát érintő megbetegedések okozta laesiokat egymástól. Előfordulhat, hogy egyazon beteg retináján a RD tünetei mellett egyéb megbetegedésekre jellemző laesiok is megtalálhatóak (4. ábra). (Mivel a cukorbetegség diagnózisa nem zárja ki egyéb betegségek előfordulását ugyanannál a betegnél.) A képfeldolgozás alapú rendszer kidolgozása során az alábbi megbetegedések jellemző retinaeltéréseinek megjelenésével kell számolnunk leggyakrabban:

- DM
- időskori macula degeneratio (AMD)
- glaucoma
- gyulladásos rendellenességek
- retinalis arteria occlusio
- naevus/melanoma



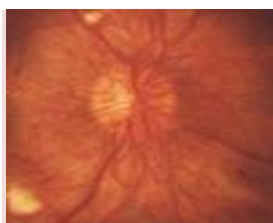
Az egészséges szemfenék képe.



Időskori macula degeneratio.



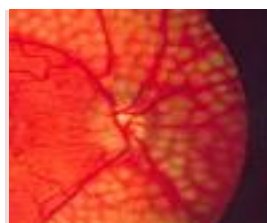
Atrophia időskori macula degeneratióban.



Diabetes mellitus szövődményeként kialakult proliferatív retinopathia (új abnormális erek és hegszövet képződése a retina felszínén).



Diabeteses macularis oedema.



A RD kezelésének eredménye (Scatter lézer sebészeti alkalmazása).

4. ábra. Az egészséges retina képe mellett látható néhány jellemző elváltozás szemfenéki képe. A szűrővizsgálatok során egyéb elváltozások felismerése nem elvárt a hagyományos szűrővizsgálatok esetén sem, így erre az automatizált eljárásokat sem kell felkészíteniünk.

2.1.4. A retinopathia diabetica okozta látásromlás

Egy az ezredforduló idején készült felmérés szerint, a világon több mint 180 millió látássérült ember élt a vizsgálat idején, akik közül 40–45 millió volt vak. A látásvesztésben szenvedők százalékos aránya Európában 0,3%. A vakok 90%-a a fejlődő országokban élt: 70%-uk Ázsiában és további 20% Afrikában. A vakok 82%-a volt 50 év alatti. A látásvesztés az esetek 80%-ában megelőzhető, vagy kezelhető lett volna [12]. Az elkövetkező húsz évben az előrejelzések szerint a vakok száma megduplázódhat. A magyarországi népességben a vakság okai között az AMD (22,7%), a RD (15,6%), a rövidlátás (13,9%), a glaucoma (12,6%), a cataracta (7,0%) és a látóideg atrophia (6,4%) fordul elő leggyakrabban. Húszéves életkor alatt a leggyakoribb vaksági ok a koraszülöttek retinopátiája, míg 40 és 60 éves kor között a rövidlátás (myopia) és a RD, 60 év felett a macula degeneratio és a RD, 80 év fölött

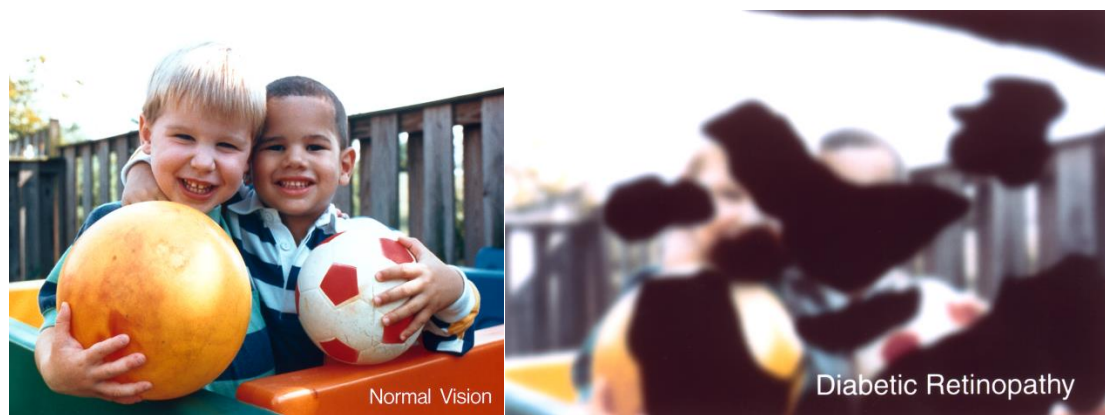
pedig a szemfenéki meszesedés (macula degeneratio) szerepel vezető okként. A cukorbetegекnél a vakság gyakorisága 25-ször nagyobb, mint az egészséges populációban [1].

Egy 1980 és 2008 között készült, 35 tanulmány adatait feldolgozó munka alapján a vizsgált 22.896 diabeteses betegre vonatkozóan a RD teljes prevalenciája 34,6%-volt (95% CI 34,5–34,8). A betegek 6,96%-a (6,87–7,04) volt a betegség proliferatív stádiumban, 6,81%-nak volt macula oedemája (6,74–6,89). A tanulmány szerint a szövödmények előfordulása erős összefüggést mutatott a DM fennállásának időtartamával, valamint a hemoglobin A1c átlagos értékével [13].

A látásromlás súlyossága, egyenes arányban áll a betegség diagnosztizálásától eltelt idővel. I. típusú cukorbetegség esetén, a betegség 15–20 éves fennállását követően, az esetek 80–95%-ában fordul elő bizonyos fokú látásromlás (RD), melynek kb. fele előrehaladott (proliferatív) forma, míg 30 év elteltével a betegek 100%-ban diagnosztizálható valamilyen fokú látóideghártya érintettség. Egy az USA-ban történt felmérés szerint az amerikai cukorbetegек 40%-ban diagnosztizálható a RD valamelyik stádiuma [14]. A II. típusú cukorbetegség esetén, a betegség 15 éves fennállását követően kb. 70–80%-ban alakul ki retinális elváltozás, melyből a diagnózis felállításának időpontjában kb. 20% esetén már megállapítható a látásromlás [14,15].

A cukorbetegség következtében kialakuló látásromlásban szenvedők száma és a betegség súlyossága függ az alapbetegség típusától, a cukorbetegség fennállásának az időtartamától és az átlagos vércukorszinttől. Mindezek mellett egyéb rizikófaktorok negatív hatásával is számolnunk kell, mint a szisztolés és diasztolés hipertónia, diszlipidémia, nephropathia és infectiók jelenléte [15].

A betegség kezdetben tüneteket nem okoz. Akár a sokáig teljesen tünetmentes megbetegedés is megjelenhet hirtelen látásromlás, vagy látásvesztés képében (5. ábra). Mivel a betegség korai felismerés esetén a súlyos szövödmények kialakulása megelőzhető, illetve a szövödmények kezelhetőek a cukorbetegség szűrése rendkívüli jelentőségű. A DM szemészeti szövödményeinek időben történő felismerésére ezért világszerte számos szűrőprogramot indítottak.



5. ábra. A bal oldali kép az egészséges látást, a jobb oldali a RD-ben szenvedő betegek által tapasztalt látásromlást érzékelteti a „beteg szemén át”.

A diabeteses betegek több mint kétharmada fejlődő országokban található (India 32 millió, Kína 21 millió), ahol az egészségügyi rendszerhez való hozzáférés nem általános [16]. A humánerőforrás magas költségigényét finanszírozni nem képes, vagy egyszerűen szakemberhiánnyal küzdő országok intenzíven keresik a jelenleginél költséghatékonyabb RD szűrő eljárásokat.

A RD kezelése önmagában is költséges, aminél még költségesebb a látóképesség elvesztésével kapcsolatos társadalombiztosítási kiadások. Ez utóbbi költségtétel is eltörpül a munkából való kiesés és az életminőség romlása miatt elszenvedett társadalmi szintű veszteségek mellett. Becslések szerint az általánosan hozzáférhető, fejlett módszertanon alapuló RD szűrés bevezetésével csak az Egyesült Államokban mintegy 167 millió dollárt és 71.000 - 85.000 visus/év-et lehetne megtakarítani [17].

2.2. A retinopathia diabetica szűrése

A RD korai felismerésének alapja a kockázatnak kitett populáció – a DM-ban szenvedő betegek – rendszeres szűrése. Az erre a célra szervezett szűrőprogramok közül az egyik legeredményesebb az „English National Screening Programme for Diabetic Retinopathy”. A jelenleg is futó szűrőprogramokban – standardok szerint képzett – szakemberek különítik el az egészséges és a RD tüneteit mutató retinaképeket, majd küldik további vizsgálatokra és kezelésre azon pácienseket, akiknél retinális elváltozást diagnosztizálnak, vagy ahol a képek értékelhetetlenek a nem megfelelő minőség miatt. A digitális képek értékelését végző szakemberek gyakran klinikai leolvasó centrumokba tömörülnek (Reading Center), ahol a számos szemészeti vizsgálóhelyről érkező digitális képeket elemzik. A vizsgálati eredményeket néhány nap elteltével küldik vissza a vizsgálat kérőinek, hogy elősegítsék az

időbeni betegellátást. A képolvasó centrumok több, akár más kontinensen elhelyezkedő vizsgálóhely összegyűjtött anyagát vizsgálják.

Hazánkban hasonló szűrőprogram beindítására voltak kísérletek, de a szűrés magas költségvonzatai miatt ezek nem váltak széles kör számára elérhetővé. Jelenleg a DM-os betegek éves szemészeti ellenőrzése ajánlott annak érdekében, hogy a betegség szemészeti szövődményeit időben felismerjék. Amennyiben a szemfenéken RD-ra utaló jel látható, a szemész szakorvos a látott kép stádiuma alapján dönt az orvos-beteg találkozás következő időpontjáról és az esetleges kezelés sürgősségéről. A központosított szűrőprogramot annak jelentős költségvonzatai miatt még nem vezették be a magyarországi szemészeti centrumokban, azonban a szakemberek nagy erőfeszítéseket tesznek a standard fotódokumentált szűrés elterjesztésére.

A RD szűrővizsgálatoknak alapvetően kétféle módszertani megközelítése létezik:

- fundus- (szemfenék) vizsgálat: direkt vagy indirekt szemtükrözéssel,
- retinafényképezés a hozzá kapcsolódó osztályozással [18,19].

Mint fentebb említettük a RD korai felismerésének alapja a veszélyeztetett betegpopuláció rendszeres szűrése. A szűrés céljára szolgáló a gyakorlatban is alkalmazott illetve a kísérleti módszerek teljesítményét az elmúlt években intenzíven vizsgálták, aminek köszönhetően viszonylag nagy mennyiségű adat áll rendelkezésünkre az egyes módszertanok értékeléséhez és egymással való összevetéséhez.

A *Health Technology Board for Scotland* megállapítása szerint a réslámpák használatával végzett indirekt szemtükrözés érzékenysége és specificitása alapján megfelelő eszköz a retinaszűrés elvégzésére. Hátránya azonban, hogy mivel a folyamat során nem készül fotódokumentáció a retináról, megoldatlan a szakemberek munkájának minőségellenőrzése és a betegség progressiójának, illetve esetenként a terápia hatásosságának nyomon követése. A direkt tükrözéses módszer szenzitivitása olyan alacsony, hogy szűrőprogramok céljára nem alkalmas [20,21].

2.2.1. A retinafényképezés

Mint láttuk, a képfeldolgozás alapú eljárások alkalmazásának előfeltétele, hogy a betegekről standard körülmények között, jó minőségű digitalizált fundus fotók készüljenek. A kétféle retinafényképezés alkalmazása megfelelő érzékenységet és specificitást biztosít ahhoz, hogy szűrővizsgálatok céljára alkalmazható legyen. A digitális fényképek készítésének előnye a képek könnyű

tárolhatósága és visszakereshetősége. Ez alapján bármikor elvégezhető a vizsgálat utólagos ellenőrzése. Nem mellékes, hogy mivel a képek elektronikus úton továbbíthatóak, segítségükkel térben elválhat egymástól a kép készítésének és vizsgálatának helye. Mindezek eredményeképpen a digitális retinafényképezés vált a szűrővizsgálatok "gold standardjává" [20]. A RD szűrésére alkalmazott standard kétmezős retinaképek példáját mutatja a 6. ábra.



6. ábra. A RD szűrésére alkalmazott standard fotódokumentáció egészséges szemfenéki képen. A kép felső sorában a jobb szemről, az alsó sorában a bal szemről készült képek láthatóak. Az első kép a maculára, míg a második a látóidegfőre centrált.

A tele-health gyakorlat

A jelenleg alkalmazott szűrőprogramokban távgyógyászati ellátási modellt (elsősorban az USA-ban) is sikerrel alkalmaznak. A modell alkalmazásának célja:

- Az ellátáshoz való jobb hozzáférés biztosítása /a RD stádiumainak meghatározásához és az adekvát kezeléshez/.
- A RD következményeként kialakuló látásélesség-vesztés előfordulási gyakoriságának csökkentése.
- A RD diagnosztikai költségeinek csökkentése.
- A távgyógyászat népszerűsítése a RD diagnosztikai-, és kezelési hatékonyságának növelése érdekében.

Azokban az országokban, ahol központosított szűrőprogramok léteznek, a programok bevezetésének indokát a megelőzhető betegség magas kezelési költségei adják.

2.2.2. A jelenlegi klinikai gyakorlatban alkalmazott szűrőeljárások bemutatása

Nagy-Britanniában készült felmérések szerint – ahol a vizsgálók betanítása és számonkérése is standardokon nyugszik – a RD detektálásának szenzitivitása 41–67% a háziorvosok, 48–82% az optometristák, 65% a szemész szakorvosok és 27–67% a diabetológusok, valamint egyéb kórházi orvosok által direkt ophthalmoscoppal végzett vizsgálatok esetén. A referálandó RD detektálásában az optometristák 77–100% közötti szenzitivitást értek el 94–100 specificitás mellett [22,23].

A fényképfelvételek elemzésén alapuló módszerek alkalmazásakor digitalizált fényképeket készítenek a betegek retinájáról, majd azokat egy későbbi időpontban az erre kiképzett szakemberek elemzik. Egy szűrővizsgálatok végzésére kiképzett szakember pupillatágítás mellett készített 45°-os retinaképek vizsgálata során 83–96%-os szenzitivitást és 83–96% specificitást tud elérni [24].

A „*The British Diabetic Association*” (Diabetes UK) által felállított standard szerint a szűrővizsgálatok szenzitivitásának el kell érnie a 80%-ot, 95%-os specificitás mellett [25,26]. A rendszeres szűrővizsgálatokat számos országban centralizálták, elsősorban költséghatékonysági és minőségbiztosítási szempontok okán [27]. A rendszerek jól üzemelnek, bár magas emberi erőforrás igényük miatt csak a fejlett gazdasággal rendelkező országok tudják finanszírozni őket [28].

A központosított RD szűrésben számos üzleti vállalkozás is lehetőséget látott, így kisebb-nagyobb arányban magán és állami tulajdonú szűrőhálózatokat is létrehoztak. Erre az egyik példa az *iScan az Inoveontól* (Inoveon Corp.; OKLAHOMA CITY, OKLA.). Az 1997-ben alapított Inoveon vállalat több mint 29 ezer klinikai iScan retinakép tanulmányozását végezte el sikeresen (7-mezős). A könnyen használható iScan-el egy technikus digitális fényképeket készít a betegek retinájáról, az elkészített képeket azután elektronikus úton küldik az Inoveon országos szűrőcentrumába, ahol képzett képelemzők értékelik a képet és alkotnak véleményt a retina állapotáról. A képek egy kiválasztott mintacsoportja ismételt vizsgálatra kerül annak ellenőrzéséhez, hogy a különböző képelemzők ugyanúgy értékelnek-e egy adott fényképet. Azok a képek, amelyek alapján nem egyértelműek a szövödményekre utaló jelek, az Inoveon szemészeti igazgatójához kerülnek továbbításra egy második szakvélemény kialakítására. Ha a képek minősége nem megfelelő az értékeléshez, az Inoveon javaslatot tesz a kezelőorvosnak a beteg beutalására egy szemészeti szakrendelésre. Az Inoveon szerint a rendszer minőségbiztosítási protokollja teljesíti, sőt túlteljesíti, a HEDIS (Healthcare Effectiveness Data and Information Set, USA) által meghatározott követelményeket a RD szűrésének vonatkozásában.

Egy másik példa a **Digiscope** (*Eli Lilly and Company; NYSE: LLY*) A Johns Hopkins University-hez tartozó Wilmer Ophthalmological Institute-nál az EyeTel által kifejlesztett DigiScope® technológia a DM-os betegeket ellátó orvosoknak segít a RD kockázatbecslésben a normál diabetológiai vizit során. A DigiScope félautomata optikai feje segítségével fényképeket készít a betegek fundusáról, ellenőrzi a képek élességét, majd az interneten keresztül elküldi a felvételeket marylandi Wilmer-EyeTel szűrőközpontjába, ahol a képek elemzése zajlik. A leleteket az orvosok 48 órán belül kapják vissza, ami alapján a betegeket szükség esetén szemész szakorvoshoz irányíthatják. A DigiScope úgy került megtervezésre és kialakításra, hogy azt bármely orvosi rendelőben, szakképzett fényképész közreműködése nélkül is lehessen használni.

Az NHS Scotland országos szinten bevezette a RD szűrését. Az internet alapú e-health technológia szállítója a Siemens Medical Solutions **Soarian® Integrated Care** (*Siemens AG; Munich; Federal Republic of Germany*). A rendszer összeköttetést teremt a betegek kezelését végző orvosok között, és lehetőséget biztosít számukra a különböző forrásokból származó képek és adatok lehívására, megtekintésére és tárolására. Az említett források a következők lehetnek: kórházi információs rendszerek (HIS), Képtároló és kommunikációs rendszerek (PACS) és digitális fényképezőgépek. Az alkalmazás lehetőséget ad jelentések elkészítésére vagy a munkafolyamatok optimalizálására is.

A **RISA** (Retinopathy Image Search and Analysis) program célja egy RD képkereső és elemző technológia "Retinopathy Image Search and Analysis" (RISA) valamint a tartalom alapú kép-visszakeresés (content-based image retrieval – CBIR) módszerének alkalmazása digitális retinaábrázoláshoz kapcsolódó gyors elemzés elvégzésére és a diagnózis kialakítására távgyógyászati modell alkalmazásával.

A központosított RD teleszűrési rendszerek általában a következőképpen működnek:

- Az orvosi rendelőben digitális fényképet készítenek a cukorbeteg retinájáról, majd
- A felvételeket elektronikus úton továbbítják a szűrőcentrumba, ahol
- Az erre kiképzett személyzet (nem orvosok) megvizsgálja a képeket a RD és a macula oedema szövödményeire utaló jeleket keresve.

A szűrővizsgálatok skálázhatóságának és költséghatékonyságának javítása érdekében számos kutatócsoport dolgozik olyan **automatizált eljárásokon**, melyek digitalizált szemfenéki fotók értékelését képesek elvégezni [29]. Az új technológiák célja, hogy bevezetésével a RD szűrése során az első vizsgálatot végző szakember, vagy szakemberek egy része kiváltható legyen. Az automatizált "előszűrést" követően csak azok a képek kerülnének a szakemberekhez, amelyek az automata szerint

kérdésesek, vagy pozitívak [30]. A kezdeti eredmények biztatóak. A szenzitivitás és specificitás értékek megközelítik a szakemberek által elért értékeket [31,32].

2.3. Az automatizált szűrést célzó eljárások bemutatása

Az automatizált eljárások fejlesztésének korai szakaszában a fejlesztések az ember által végzett szűrést próbálták utánozni automatizálható módszerek segítségével és egyértelműen a humán erőforrás kiváltását tűzték ki célul. A korai tapasztalatok ezt követően nyilvánvalóvá tették, hogy a különböző projektek során kifejlesztett rendszerek erre nem lesznek alkalmasak, viszont a betegek előszűrése automatikus rendszerek segítségével reális célkitűzés lehet. Utóbbi esetben a szakembereknek csak a pozitívnak, kétesnek, vagy értékelhetetlennek minősített felvételeket kell megvizsgálniuk. Ennek eredményeként jelentősen csökkenhet a szűrővizsgálatok emberi erőforrás igénye, ezáltal költsége is. A költségek csökkenésével olyan helyeken is szervezhetőek szűrőprogramok, ahol eddig azokat nem tudták finanszírozni. A fejlesztések túlnyomó többsége a digitális retinaképek képfeldolgozási módszerek segítségével történő automatizált elemzésén és osztályozásán alapszik.

2.3.1. A képfeldolgozás alapú módszerek általában

A képfeldolgozás alapú módszerek a szemfenékről készített digitális fotón számítógépes algoritmusok segítségével ismerik fel azokat az elváltozásokat, amelyek a RD-ra jellemzőek. Az egyik ilyen szemfenéki elváltozás, amelynek detektálására a napjainkban elterjedt algoritmusok többsége épül a MA. A detektálás során az algoritmusokat alkalmazó szoftver azonosítja azokat a szemfenéki laesioakat, amelyek feltételezhetően MA-k. Ezt követően a képeket osztályozza annak megfelelően, hogy melyeknél nagyobb és melyeknél kisebb a valószínűsége, hogy a képen egy vagy több valódi MA található. Az előbbi eseteket a "beteg - RD" utóbbiakat a "nem beteg – non-RD" kategóriába sorolja.

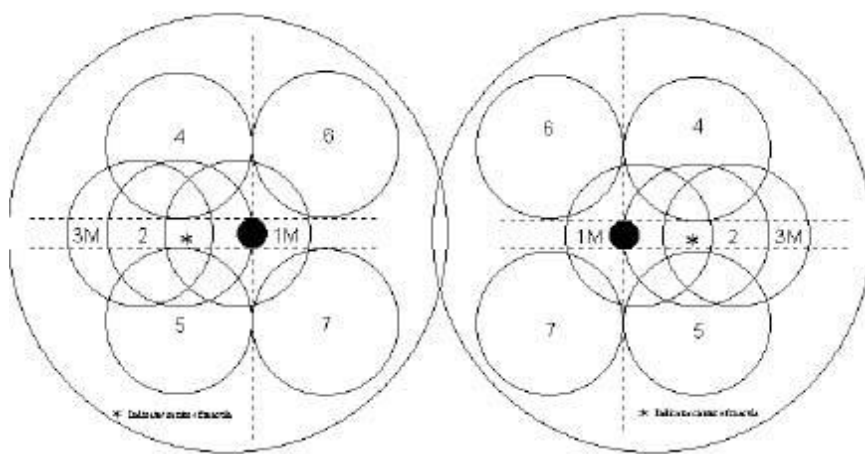
Ha az automatizált rendszereket a jövőben a klinikai rutinban szeretnénk vizsgálni, akkor azoknak alkalmazkodniuk kell a betegellátás során jelenleg alkalmazott protokollokhoz. A fejlesztés alatt álló rendszerek célja a RD szűrése, azonban figyelembe kell vennünk, hogy a szemfenéken más megbetegedések is jellemző eltéréseket okozhatnak. Az automatizált képfeldolgozás alapú eljárásoknak működésük során a RD-ra jellemző elváltozásokat más megbetegedésre jellemző eltérések jelenléte mellett is fel kell ismerniük, hiszen a cukorbetegség nem zárja ki más retinát érintő kórképek párhuzamos előfordulását. Az automatizált rendszereknek ezeket az egyéb betegségekre jellemző elváltozásokat el kell tudniuk különíteni az egészséges retina képétől, azonban nem szükséges az egyes laesioak azonosítása.

2.3.1.1. A bemeneti, digitális kép létrehozása

A rendszerek inputját digitális képek jelentik. A klinikai gyakorlatban egy-, két-, három-, és hétmezős digitális képek készülnek. A mezők száma megegyezik az egy retináról készített képek számával. Az egyes képek a retina különböző részeit ábrázolják egymással átfedésben.

Az *Airlie House* hétmezős osztályozórendszerben (7. ábra) minden szemfenékről hét fénykép készül különböző nézőpontokból. Ez lehetővé teszi a képek későbbi háromdimenziós rekonstrukcióját is. (A retina képletei és az elváltozások gyakran jelentősen előemelkednek a retina síkjából, s bizonyos elváltozások felismerését (pl. macularis oedema) a 3D képek analízise jelentősen megkönnyítheti.

(<http://www.abdn.ac.uk/~opt065/Web%20pages%20only/drh-16.htm>)



7. ábra. Az *Airlie House* osztályozórendszer.

Az inputot jelentő képek szabványosítása során a kezdeti években a rendszerek fejlesztőknek gondot okozott, hogy az akkoriban széles körben készített fekete-fehér, vagy papír alapú hagyományos fényképek digitalizált változatait is feldolgozzák-e. Később a rendszereket az egy illetve kétmezős digitális retinafotókra optimalizálták, ami aztán megkönnyítette a több klinikai centrumra kiterjesztett validálást, valamint a fejlesztői csoportok által elért eredmények nemzetközi összevetését, a munkacsoportok közötti versenyek szervezését.

2.3.1.2. A képek színtartományának és normalizálási módjának kiválasztása

A képfeldolgozás alapú szűrőrendszer kidolgozása során figyelembe kell vennünk azt a tényt, hogy egyes laesiók jobban láthatóak a vörös szín kiszűrésével készült képeken. A színcsatornák megválasztása mellett fontos előfeldolgozó lépés a szemfenéki képek normalizálása (standardizálása) azok méretére, intenzitás eloszlására, zajcsökkentő elmosására vonatkozóan.

2.3.1.3. Számítógéppel támogatott képanalízis

Az értekezésben bemutatott munka előkísérletei során a normál szemfenéki struktúrák (pl.: erek, látóidegfő, sárgafolt) azonosítását lehetővé tevő algoritmusokat fejlesztettünk ki. Ezt követően a kóros retinára jellemző morfológiai elváltozásokat megbetegedések szerinti csoportokba soroltuk. Az egyes diabetesre jellemző laesiok felismeréséhez specifikus algoritmusokat hoztunk létre, majd külön algoritmusokat dolgoztunk ki a szemfenék durva (nem DM specifikus) elváltozásainak előszűrésére. Az értekezésben bemutatott képfeldolgozás alapú eljárásban ezek közül kizárólag az előfeldolgozási lépéseket és a MA detektálására szolgáló algoritmusokat alkalmazzuk.

2.3.1.4. A normál szemfenéki struktúrák és a laesiok felismerése

A munkánk előkísérletei során kifejlesztett döntéstámogató rendszerek az egészséges retina jellemzőit, valamint az említett kóros laesiokat is objektumokként, vagy objektumok csoportjaként azonosították. Az objektumok azonosítása után azokat egészséges és beteg (RD) csoportba sorolták. Jelen munka keretében megfelelő eredménynek tekintettük pusztán a szűrés elvégzését és a retina elváltozásai közül kizárólag a MA-k azonosítására fókuszáltunk. Ennek oka az, hogy a hasonló, csak MA-kat detektáló megoldásokról áll rendelkezésünkre a legtöbb adat és ezek a megoldások állnak legközelebb a rutin klinikai alkalmazáshoz. Emellett szerettük volna, ha a saját megoldásunk a későbbiekben könnyen összevethető a más munkacsoportok által fejlesztett rendszerek teljesítményével.

A retina laesiok közül néhány több betegségre is jellemző lehet. Más laesiok kizárólag egy meghatározott rendellenességhez kapcsolódnak, vagy jellemző kombinációk részeként jelennek meg. Például a Drusen és a kemény exudatum egyformán sárgásfehér foltként jelenik meg a kétdimenziós képeken. A spektrál-analízis segítségével a két elváltozás elkülöníthető egymástól. Ezek az elváltozások méret-, szín- és alaktani sajátosságaikon alapuló, valamint a körülöttük elhelyezkedő addicionális képletek (pl. exudatum esetén kapcsolódó MA) tanulóalgoritmusokkal megkülönböztethetők egymástól.

2.3.1.5. Tanulóalgoritmusok alkalmazása

A retina egyes elváltozásainak azonosítására kidolgozott legmodernebb eljárások tanítóalgoritmusokra építenek. A leggyakoribbak az apró, vörös elváltozások vagy a világos (fehéres/sárgás) elváltozások detektálására és osztályozására kidolgozott eljárások [33,34]. A betanításhoz a képelemekhez (például képpontokhoz vagy nagyobb képi egységekhez) sajátságvektorokat rendelnek, ahol az egyes sajátságokat (skalár értékeket) többnyire szélesebb körben is használt sajátságkinyerőkkel határozzák meg.

Többváltozós statisztikai módszerekkel (például faktor/regresszió analízissel) ellenőrizhetjük az egyes sajátságok redundanciáját, relevanciáját, illetve olyan tanuló- és osztályozó algoritmusokat fejleszthetünk, amelyek figyelmen kívül tudják hagyni a kevésbé releváns sajátságokat. Az egyik ilyen korszerű osztályozó algoritmus például a support vector machine (SVM). A SVM-et az értekezésben bemutatott munka során is használtuk.

2.3.1.6. A retinopathia diabetica szűrésére alkalmas képfeldolgozó rendszer működése

A kép háttérének eltávolítása - Maszkgenerálás

A retináról készített felvételeken a retina képleteinek adatait el kell választanunk a kép háttérétől. Ekkor a minket érdeklő területeket külön vizsgálhatjuk, kiküszöbölve a háttér zavaró hatását.

A retina képére félkör alakú maszk helyezhető, amelyet a zöld színcsatorna intenzitásának változtatásával hozunk létre. Kissé precízebb módszert kapunk, ha mindhárom színcsatorna (RGB) intenzitását változtatjuk. A képen található nem kívánatos artefactumok megfelelő matematikai morfológiai operátorok (erózió, dilatació, nyitás, zárás) alkalmazásával könnyen eltávolíthatóak.

A digitális kép minőségének javítása

Retinafényképezés közben számos olyan zavaró tényező léphet fel, mely rontja a képminőséget (pl.: cataracta, pislogás, nem megfelelő fényviszonyok, homályos elülső szegmentum, cornealis elváltozások), ezért a bevitt adatok minőségének ellenőrzése és javítása fontos eleme az automatizált döntéstámogató rendszer fejlesztésének. (Például a fentebb említett Eye Check program esetében a készített felvételek 12%-a szemészetileg nem volt értékelhető.

További problémát okoz a képek értékelése során a nem megfelelő kontraszt, ez különösen a fovea pontos helyének meghatározásakor jelent problémát. A képélesség javítása mellett szükség van a színek normalizálására is.

A szemfenéki kép normál képleteinek azonosítása

Miután a fent leírtaknak megfelelően megtörtént a kép minőségének javítása, a normál szemfenéki képletek azonosítása következik. Az azonosított képletek normál és kóros voltának elkülönítésére laesio specifikus algoritmusokat fejlesztünk.

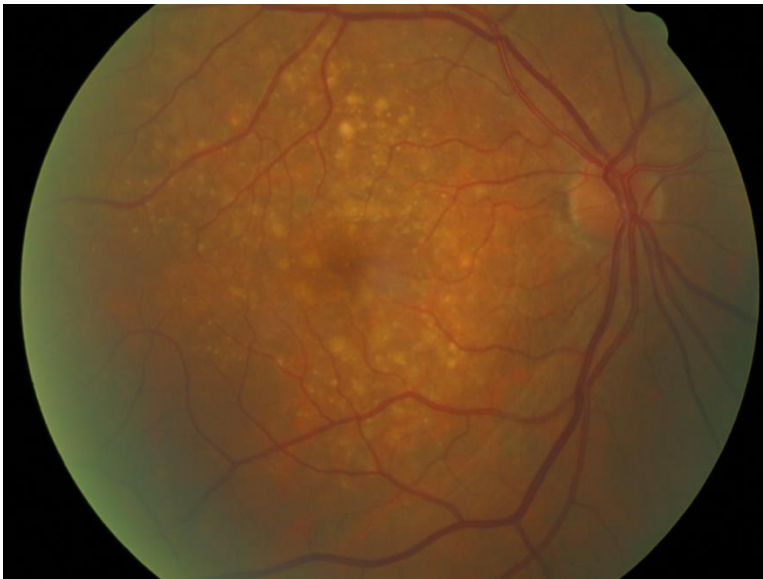
Az érhálózat a szemfenék egyik legmegbízhatóbban detektálható anatómiai képlete. Az alaposan feltérképezett érhálózat segítségével más képletek (látóidegfő, sárgafolt) helyzete is könnyebben meghatározható. A szakirodalomban számos technikát leírtak az érhálózat szegmentálására vonatkozóan. A laesio több típusa is csak megfelelően feltérképezett érhálózat segítségével azonosítható.

A kép orientációja pontosan meghatározható a helyi tájékozódási szisztéma szerint, amelynek a kiindulási pontja a sárgafolt centrális része (fovea), horizontális tengelye pedig a látóidegfő irányába mutat. Egészséges retina esetében a látóidegfő könnyedén azonosítható világos foltként jelentkezik a retina ereinek találkozási pontjánál.

A fovea azonosítása a retina feltérképezésének következő lépése. A látóidegfő és a fovea közötti távolság közel állandónak tekinthető. Ezen geometriai sajátosság figyelembe vételén kívül a fovea helyének meghatározásakor segítséget jelenthet a terület alacsonyabb fényintenzitása is.

A leggyakoribb retina laesiok detektálása

A retina különböző rendellenességei különböző vizuális sajátságokként jelennek meg a szemfenékről készített képeken (8. ábra). Jellemezhetőek például a színük, alakjuk és méretük alapján. A kemény exudatumok, gyapottépésszerű gócok és Drusenek világos foltokként jelennek meg a szemfenéki képeken.



8. ábra. A képen az időskori népességben viszonylag gyakran előforduló elváltozások láthatóak: kemény és puha Drusenek az éleslátás helyén, mely nem RD-ra jellemző elváltozás.

A kemény exudatumok fehér vagy sárga foltjainak kimutatása lehetséges például a RGB színtartomány megfelelő lineáris transzformációjával. A gyapottépésszerű gócok sárgásfehér foltokként jelentkeznek a retinán. A Drusen ugyan az AMD-re jellemző elsősorban, de idősebb életkorban egészséges személyeknél is előfordul, kis-méretű fehér, vagy sárga elváltozás képét mutatva. Nagyobb méretű vagy számú Drusen jelenléte a maculában azonban kóros állapotot jelez (AMD). Az osztályozás

során az algoritmusok kifejlesztését nehezíti, hogy alkalmasaknak kell lenniük a RD-ra jellemző elváltozásoknak a Drusenektől való elkülönítésére, viszont az egyes elváltozások egyértelmű azonosítására a szűrés során nincs szükség, így 2D képeken a gépi tanulás jól alkalmazható eszköz erre a célra.

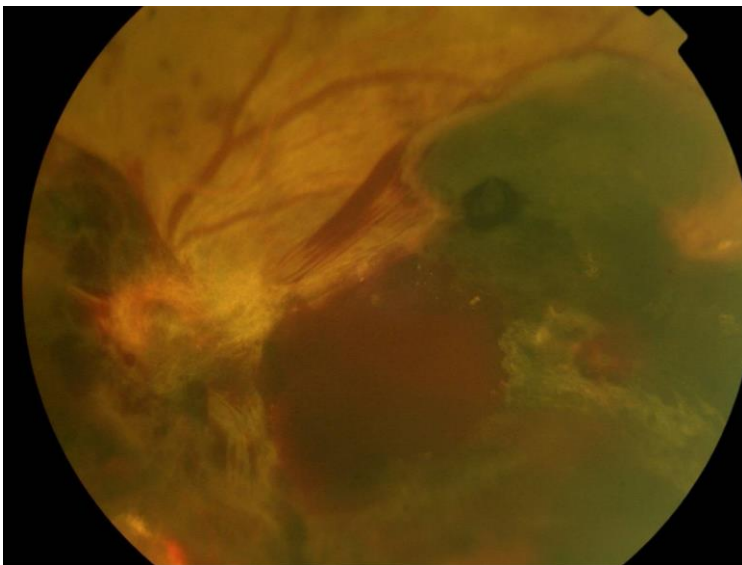
A **vérzések és MA-k** vörös laesioként jelentkeznek a szemfenéki képeken. A MA-k apró, kerek, vörös elváltozások, míg a vérzések lehetnek pont, folt és láng alakúak. A színük hasonló a véredényekéhez, ezért az azonosításuk érdekében fontos a laesiok kontúrjainak élesítése.



9. ábra. Microaneurysmát ábrázoló kép, mely a RD kialakulásának korai jele.



10. ábra. A RD előrehaladtával gyapottépeses gócok és egyre kifejezettebb vérzések jelennek meg a szemfenéken.



11. ábra. Proliferatív RD képe, látóideghártya és üvegtesti vérzéssel.

A **vénás hurkok** detektálásának feltétele a szemfenék érhálózatának pontos feltérképezése. Matematikai jellemzésük például egy absztrakt gráf keringési rendszerhez való kapcsolásával lehetséges, ahol a véredények vastagsága pixelben kifejezhető. Detektálásukban fejlett matematikai morfológiai eszközök (például vázkijelölés) használható.

A **macularis oedémára** a 2 dimenziós (D) képen az alábbiak utalhatnak: i.) a foveolától 1 látóidegfő átmérőn belül fellelhető kemény exudatumok; ii.) csoportban előforduló exudatumok a makulában; iii.) a foveolától 1 látóidegfő átmérőn belül fellelhető MA vagy vérzés, melyet 3D megjelenítés esetén a foveolától 1 látóidegfő átmérőn belül fellelhető retinális megvastagodás egészít ki.

A RD diagnosztizálása során a 3D alapú képanalízis kiváltható 2D alapú képanalízissel. A 3D alapú megjelenítés pontosítja ugyan a 2D alapú képmegjelenítést, de az oedema egyértelmű diagnózisa, valamint az elvégzendő terápiás lézerkezelés megtervezése további fluorescein angiographiás vizsgálatot tesz szükségessé. Saját munkánk során ezt az eljárást nem alkalmaztuk, mivel a RD automatizált szűrése során jelentősége elhanyagolható.

Döntéshozatal

A laesiok azonosítása után a képeket kóros és egészséges csoportokba szükséges sorolni. Az osztályozás történhet például neurális hálózatok, SVM, vagy Bayes-hálók segítségével. Két utóbbit munkánk során alkalmaztuk is, így alkalmazásukat a módszertani fejezetben részletesen bemutatjuk.

A szűrőrendszer működésének lépései

A fentieket összegezve tehát általánosságban egy automatizált szűrő rendszer működése a következő lépésekből áll:

- Maszkgenerálás (az érdeklődési terület elválasztása a háttértől)
- Képmínőség javítása (szín/fényminőség normalizálásával)
- Digitális kép jellemzőinek kinyerése (beleértve az egészséges képleteket és számos laesiot, a mi esetünkben MA-kat)
- Döntéshozatal.

2.3.2. Microaneurysma detektor – a leggyakrabban alkalmazott képfeldolgozás alapú szűrőeljárás

A retinát érintő laesiók nagy száma és variábilis megjelenése miatt az automata rendszerek fejlesztése során kifejlesztett algoritmusok olyan elváltozás detektálására fókuszálnak, mely könnyen leírható, viszonylag konzekvens megjelenésű és érzékeny, korai indikátora a RD kialakulásának. Ez az elváltozás a MA, mely a RD legkorábbi tüneteinek egyike, és jelenléte nélkül a RD diagnózis nem állítható fel [9, 10, 11]. Ennek tükrében nem meglepő, hogy az automatizált szűrőeljárások többsége kizárólag ezen elváltozások kimutatására koncentrál, figyelmen kívül hagyva az összes többi elváltozást.

A fejlesztések eredményeként a MA detektorok specificitási és szenzitivitási értékei közelítik a humán vizsgálók eredményeit. Az automata vizsgálatok nagy előnyének tartják, hogy a szoftverek egy adott vizsgálohelyen és betegpopuláción stabilan hozzák a beállításkor mért értékeket, míg a humán vizsgálók eredményei egymástól illetve napszaktól és egyéb tényezőktől függően eltérnek [35,36,37].

Az egyes fejlesztői csoportok eredményeinek összehasonlíthatóvá tételének és összehasonlításának céljával a University of Iowa MA detektor fejlesztői versenyt szervezett (International Retinopathy Online Challenge), melyen munkacsoportunk tagjai 2010-ben a legjobban teljesítő algoritmus díját szerezték meg [38].

Ha egy képfeldolgozás alapú döntéshozó rendszernek egyféle elváltozás felismerésére kell fókuszálnia, akkor az eredmények lényegesen jobbak és reprodukálhatóbbak.



12. ábra. A szabad szemmel alig észrevehető MA-k, hajszálvékony erek kis buborékszerű kiöblösödései. A színes szemfenéki képen a MA-k piros pontok képében jelennek meg.

Az automata előszűrés működési elve az *igen/nem* (RD/non-RD) döntéshozó mechanizmuson alapszik. Durva közelítéssel azt mondhatjuk, hogy ha a képfeldolgozó eljárás alkalmazása során a szemfenéki képen az algoritmus nagy biztonsággal MA-t detektál, a vizsgált képet a RD csoportba sorolja. Ezen csoportba kerülnek azok a képek is, melyeket az algoritmus valamilyen ok miatt nem tud besorolni az egészséges csoportba, azaz „kétes”-nek ítél meg.

Ennél a módszernél gyakrabban alkalmazzák azt a technikát, amikor a képeket gépi tanuló algoritmusokat alkalmazva a képfeldolgozó algoritmusok által azonosított MA-k száma alapján sorolják be a képet RD, vagy non-RD csoportba. A RD csoportba sorolt képeket a szűrőrendszer szakember általi vizsgálatára továbbítja.

Az ellátórendszer szempontjából a nyereség az, hogy a ”biztosan” egészséges kategóriába sorolt képeket szakembereknek nem kell látnia. Ezzel nagy populációk szűrése esetén jelentős költség és idő takarítható meg.

2.3.3. Könnyproteomikai alapú eljárások

A könnyfolyadék összetétele igen komplex nagy mennyiségű fehérje mellett jelentős mennyiségben tartalmaz lipideket, sókat, mucint és egyéb szerves molekulákat. Elsődleges szerepe a

cornea nedvesen tartása és síkosítása, valamint első védelmi vonalként szolgál a kórokozók elleni védelemben [39,40]. A könny összetevői közül a fehérjék fordulnak elő a legnagyobb mennyiségben, hozzávetőlegesen 95%-át adva a könnyminták száraz tömegének [41]. A fehérjék elsődleges forrásai a könnymirigyek. Emellett a szem felszínét borító epithelből és a vérből származó fehérjék szintén jelentősebb mennyiségben fordulnak elő. Több mint 500 azoknak a fehérjéknek a száma, amelyek igazoltan kimutathatóak a könnyfolyadékban. [42,43].

A könny fehérjeösszetétel változások elemzésével több kutatócsoport is foglalkozott az elmúlt években, akik részletesen leírták a fehérjeösszetétel megváltozását például a sebgyógyulás illetve gyulladásos megbetegedések kapcsán. Elektroforézises és tömegspektrometriás technikákkal végzett vizsgálatok támasztják alá azt a tényt, hogy az egészséges és DM-ban szenvedő személyek könny fehérje összetétele jellemző eltéréseket mutat [39, 40].

A könnyfolyadék fehérjeösszetételének vizsgálatára olyan kvantitatív és kvalitatív technikákat is alkalmaznak mint az egy és kétdimenziós elektroforézis, az ELISA és Western blot [44,45,46]. Az utóbbi időkben nagy szenzitivitású és nagy áteresztőképességű módszereket kezdtek alkalmazni a megbetegedések és traumák fehérjeösszetételre gyakorolt hatásának vizsgálatára. A MALDI-TOF, SELDI-TOF és a folyadék kromatográfiával kapcsolt LC/MS tömegspektrometriás analízisek segítségével a könnyminták mikroliteres mennyiségeiből is elérhetővé vált a fehérje profil meghatározása [47,48,49,50]. Ismert tény, hogy RD fennállása esetén a szemben lévő erek falának permeabilitása megváltozik, ezért azt feltételeztük, hogy ez hatással van a könny fehérjeösszetételére is [51, 42].

Egyes kutatócsoportok hasonló jelenséget figyeltek meg a szem belsejét kitöltő üvegtestnél. Azonosítottak olyan fehérjéket, melyeknek az üvegtestben mért koncentrációja RD esetén megemelkedik [52].

A retina kóros állapotai az üvegtest fehérjeösszetételére is hatnak különösen akkor, ha a BRB sérült [53]. A BRB sérülése gyakran előfordul RD, AMD, retina véna okklúzió, uveitis és egyéb krónikus retinát érintő megbetegedések esetén [54]. A RD komplex rendszerbetegség, amely érinti a kis és nagyerek falait, gyulladásos folyamatokat az idegsejtek elhalását idézheti elő hő-sokk proteinek (pl. crystallin) fehérjék termelődését váltva ki ezzel. A sérült BRB funkció központi mechanizmusa a retina endothel sejtek megváltozott permeabilitása, melyet a növekedési faktorok, citokinek és advanced glycation end products (AGEs) megemelkedett szintje, a gyulladásos folyamatok és a pericyták elvesztése. Ezzel párhuzamosan az endotheliális sejtek közötti kapcsolatok megnyílásával, valamint a caveoláris transzport mechanizmusok sérülésének mennyiségi és minőségi változásával fokozottá válnak a sejteken keresztüli, illetve sejtek közötti transzport folyamatok.

A pericyták és astrocyták funkcionális változásai, az endotheliális glycocalyx és az endotheliális lamina basalis strukturális változása szintén hozzájárulnak a BRB áteresztőképességének növekedéséhez. A Starling-mechanizmusnak megfelelően a plazma proteinek BRB endothel sejtek okozta fokozott transzcelluláris transzportja megnöveli az interstitiális tér ozmózis nyomását, amely az egyik legmeghatározóbb faktor a macularis oedema kialakulása során [54].

Az újabb teljes fehérje analízist lehetővé tevő vizsgálatok azt mutatják, hogy az általános stressz-válasz a hő-sokk fehérjék szintjének emelkedéséhez vezet. A retinában expresszált alpha-crystallin DM-ban overexpresszálttá válik, mely a retina sejtek adaptív válaszában tekinthető [55]. Ezek a változások lényegesen specifikusabbak RD-re, mint a könny proteom változásai, azonban az üvegtestből történő mintavétel invazív módja jelentősen korlátozza a módszer alkalmazhatóságát. Mivel a crystallin számos kórfolyamatban is emelkedett lehet, biomarkerként való használhatóságát alacsony specificitása is gátolja.

RD biomarkerek

Az utóbbi néhány évben számos potenciális biomarkert azonosítottak immunhisztokémiai, qPCR alapú mRNS-expressziós vizsgálatok vagy individuális fehérje koncentráció mérés segítségével. Ilyenek voltak például a szérumban lévő retinol kötő protein 4, a plazma apelin, az üvegtestben lévő PDGF és VEGF, a könny lipopoliszacharid (LPS)-kötő protein és szolubilis CD14, melyeknek emelkedett szintjét mérték proliferatív RD-ban szenvedő betegeknél [56,57,58,59,60]. A projektek az egyes fehérjék célzott analízisén alapultak, így nem adtak rálátást a fehérjeösszetétel globális mintázatának változásaira. Egyes kutatócsoportok vizsgálták a neovascularizált membrán és a nem vaszkuláris epiretinális membrán összetételét proliferatív RD-ban szenvedő betegeknél poolozott mintákon LC-MS/MS metódust használva. A vizsgálatok a periostin és a pigment epithelium derived factor (PEDF) szintjét találták szignifikánsan magasabbnak a RD-ás betegeknél [61].

Mások RD vérplazma protein biomarker panelt hoztak létre kvantitatív tömegspektrometria alkalmazásával [62]. Az eredmények hasznosíthatóságában erős korlátot jelent az a körülmény, hogy a vérben lévő fehérjék koncentrációváltozásai sokkal inkább a DM kórjelzői és kevésbé korrelálnak a szemben bekövetkező elváltozásokkal.

Munkacsoportunk volt az első, aki a RD hatására bekövetkező könny proteomikai mintázat változását vizsgálta gépi tanulás alapú módszertannal. Könnyminták noninvazív módon gyűjthetők, emellett a könny fehérjeösszetételének változása várhatóan jobban korrelál a szemben bekövetkező változásokkal, mint a perifériás vér fehérjeösszetételének változása. Várakozásaink szerint ez a két tényező teheti a könny proteomikai megközelítést alkalmassá jövőbeni klinikai diagnosztikai felhasználásra.

Az általunk is alkalmazott iTRAQ technológiát korábban sikerrel alkalmazták száraz szem szindróma biomarkerek azonosítására [63]. Az iTRAQ kvantitatív proteomikai eljárás egy robosztus technika, mely lehetővé teszi a fehérjék kémiai jelölését különböző stabil izotópot tartalmazó kódolt markerekkel, így több minta egyidejű vizsgálatát a kísérletek során [64,65].

Munkacsoportunk jelen munkát megelőző előkísérletei során a RD betegektől és a kontrolloktól származó iTRAQ jelölt könnymintákat nanoHPLC vel kapcsolt nano-ESI MS/MS 4000 QTRAP tömeg spektrométer segítségével vizsgálta és megkísérelte RD-ra jellemző fehérje biomarkerek azonosítását. Ezen kísérletek elsődleges célja a kvantitatív proteomikai technikák alkalmazása volt a RD progressziójának monitorozására. A vizsgálat során ennek megfelelően olyan biomarkereket kíséreltünk meg azonosítani, melyek a RD proliferatív stádiumának kialakulását képesek előrejelezni olyan korán, amilyen korán csak lehetséges. 6 potenciális fehérje biomarkert azonosítottunk a vizsgálatok során: lipophilin A, lipocalin 1, lactotransferrin, lacritin, lysozyme C és immunoglobulin lambda lánc [51].

A legújabb gépi tanulás alapú módszerek képesek arra, hogy a betegek könnymintáiban megtalálható összes fehérje koncentrációjának „globális mintázatát” figyelembe véve végezzék el az elemzéseket és a vizsgált személy besorolását (RD/non-RD) csoportokba [66].

A diagnosztikai panel által alkalmazott markerek számát a klinikai rutinban való könnyebb alkalmazhatóság érdekében a jövőben csökkentenünk kell, ami lehetővé tenné, a kísérletek során alkalmazott és még relatíve költséges tömegspektrometriai technikák kiváltását olcsó, gyors és könnyen használható technológiával.

2.3.4. Kombinált eljárás – a könnyproteomika és a képfeldolgozás kombinálása

Hasonló eljárás leírását a tudományos szakirodalomban nem találtunk. Munkacsoportunk legújabb kutatásai a könnyproteomikai vizsgálatok és a képfeldolgozás alapú vizsgálatok kombinált alkalmazására irányulnak. A koncepció hátterét az a feltételezés adja, hogy eltérő technológián és módszertanon alapuló független, eltérő adatforrásból származó bemeneti adatok együttes vizsgálata javíthatja a szűrőeljárás pontosságát. Elképzelésünk szerint a módszertan a klinikai rutinban úgy működne, hogy a szemfenék vizsgálat (fotózás) időpontjában egy könnyfehérje marker gyorstesztet is elvégeznének. A képfeldolgozás és a könnyfehérje teszt eredményét egy orvosi döntéstámogató szoftver együttesen értékelné. A kombinált vizsgálat eredménye vélhetően megbízhatóbb eredményt biztosítana, mint az önállóan alkalmazott képfeldolgozó, vagy könnyproteomikai eljárás. Munkánk második fázisában ezt a kérdést tanulmányoztuk.

2.4. Célkitűzések

Munkacsoportunk 2008-óta foglalkozik képfeldolgozás alapú módszertanok kidolgozásával. A képfeldolgozás alapú eljárások szakirodalma rendkívül szerteágazó. Saját korai kísérleteink ezen eljárások eredményeinek reprodukálására, később fejlettebb eljárások kidolgozására irányultak.

A képfeldolgozási módszertan alkalmazása a DM-ban szenvedő betegek fundusáról készített digitális képeken történik. Az elemzés során azonosítani kívánjuk azokat a szemfenéki elváltozásokat, amelyek megjelenése a RD-ra jellemző. Az általunk kidolgozott eljárások egy része több RD-ára típusos elváltozásféleség egyidejű felismerésére volt képes. Más eljárásokat kizárólag egy elváltozás, a MA felismerésére hoztunk létre. Az értekezés alapjául szolgáló munka 2. fázisa során egy mások által kifejlesztett MA detektort implementáltunk és használtunk célkitűzéseink megvalósításához (ahogy azt a 3.2.2. fejezetben részletesen bemutatjuk).

Az értekezés alapjául szolgáló munka 1. fázisában egy a területen eddig nem alkalmazott módszertanban, a könny-proteomikai vizsgálatokon alapuló RD szűrésben rejlő lehetőségeket vizsgáltuk. Munkánk háttérét az a felismerés adta, hogy a különböző szemészeti kórképekben megváltozik a könnyfolyadék fehérje összetétele [39, 40]. A betegségek hatására kialakuló könnyfehérje összetétel változás mintázata jellegzetes lehet bizonyos megbetegedésekre, ami alkalmassá teszi őket szűrő és diagnosztikai eljárásokban való alkalmazásra. Feltételeztük, hogy a RD esetében is találhatunk ilyen jellemző fehérje koncentráció mintázatbeli eltéréseket.

Jelen munka során a fehérje koncentrációk globális mintázat alkalmazásával megkíséreltük értékelni a módszertan potenciális teljesítményét, valamint a markerszám csökkentésének hatását a vizsgálat szenzitivitás és specificitás értékeire. Ezt követően (munka 2. fázisa) kombináltuk a könnyproteomikai és a képfeldolgozás alapú módszertanokat. *Értekezésemben a könnyproteomikai illetve a kombinált módszertan vizsgálatával kapott eredményeket mutatom be.*

Az 1. munkafázisában a könnyben található fehérjék RD-markerként való alkalmazhatóságának vizsgálatára fókuszáltunk. Előkísérleteink során sikerrel azonosítottunk olyan fehérjéket, melyek a RD-ban bekövetkező jellemző koncentrációváltozásuknak köszönhetően markerként használhatóak. Munkánk célkitűzései a következők voltak:

Célkitűzés I.: A könny fehérjék koncentrációjának egyedi vizsgálatát követően nagyszámú fehérje koncentrációjának egyidejű meghatározásával olyan globális mintázatokat kívántunk azonosítani, melyek a RD-ra jellemzőek. Feltételezésünk az volt, hogy ha egyidejűleg több fehérje RD okozta változását is figyelembe vehetjük, a szűrt populáció precízebb osztályozására leszünk képesek.

Célkitűzés II.: A könnyfehérje koncentrációk globális mintázatának vizsgálatát követően szeretnénk volna a vizsgált fehérjék számát csökkenteni úgy, hogy azzal a módszer diagnosztikai pontosságát ne csökkentsük jelentősen. Az alacsony prediktív értékkel bíró markerek kizárása, vagyis a vizsgált markerek számának csökkentése, az eljárás költséghatékonyságát növelni, kivitelezhetőségét pedig egyszerűbbé, gyorsabbá tenni volt hivatott. Elképzelésünk szerint ezzel az eljárás klinikai rutinban való alkalmazásának esélyét növelhetjük. Távlati célunk olyan a járóbetegellátásban is használható gyorsított kidolgozása, mely a könnyből képes rövid idő alatt a RD jelenlétét megállapítani.

A 2. munkafázisban a könnyproteomikai és a képfeldolgozási módszertanok kombinálásán alapuló RD szűrő eljárást dolgoztunk ki és vizsgáltunk. Ennek kapcsán az alábbi célkitűzést fogalmaztuk meg:

Célkitűzés III.: Vizsgálataink során a szakirodalomban leírt eredményekkel való összehasonlíthatóság érdekében egy a mások által kidolgozott és publikált képfeldolgozás alapú eljárást implementáltunk. Az eljárás a RD korai jeleként számon tartott MA-k detektálásán alapult. Miután az eljárással kapott eredményeket reprodukáltuk, bizonyítani akartuk, hogy a fentiekben leírt könnyproteomikai eljárással kombinálva a kombinált módszertan teljesítménye meghaladja i.) a fent leírt tisztán könnyproteomikai markerek használatára épülő eljárás teljesítményét ii.) és a tisztán képfeldolgozás alapú módszertan eredményét is. A kombinált eljárás használata során a képeken azonosított MA-k számát és a könnyfehérje koncentráció értékeit, mint a betegség jellemzőit együttesen vettük figyelembe.

3. Metodikák

A metodikák bemutatásánál külön mutatjuk be a könnyproteomikai markerek RD szűrésére való felhasználhatóságának vizsgálatára irányuló kísérleteink során használt módszertant illetve a kombinált, képfeldolgozási és könnyproteomikai markereket együttesen használó kutatásainkhoz kapcsolódó eszközöket és módszertant.

3.1. A könnyproteomikai alapú szűrés módszertanának vizsgálata során alkalmazott módszerek

3.1.1. Betegvizsgálat és mintavétel

A munka első szakaszában 119 diabeteses beteg került bevonásra a Debreceni Egyetem Szemklinikájának járóbeteg szakrendeléséről. 73 betegnél csak az egyik szemből vettünk könnymintát, a kizárási kritériumok teljesülése miatt, melyek az alábbiak voltak: szemfelszíni betegség (pl.:

keratoconjunctivitis sicca), operált szem, nem diabeteses eredetű szemfenéki elváltozás, illetve a beteg együttműködési hajlandóságának hiánya. A vizsgálatból azon szemeket is kizártuk, ahol a retinafotózás nem volt kivitelezhető, mivel a RD diagnózisának megalkotásához digitális szemfenéki képeket használtunk. A retinafotók elkészítésének akadályát jelentették a cataracta, vérzés, zárt zugú glaukóma, vagy technikai nehézségek. A RD diagnózisát 7-mezős fundus fotográfiák két független szemész általi vizsgálatával kaptuk. A fotók készítéséhez Megaplug kamerát használtunk: Model 1,6i/10 BIT Zeiss (Carl Zeiss Ophthalmic System A6, Jena, Germany). A 165 vizsgált szem közül 55 volt egészséges és 110 mutatta a RD jeleit. Egy beteg esetén csak a beteg egyik szeme mutatta a RD jeleit.

A könny minták gyűjtését arra kiképzett asszisztens végezte standardizált protokollt követve [67]. A mintavétel üvegapilláris segítségével történt közvetlenül a fundus vizsgálat során végzett pupillatágítást megelőzően, a laterális canthusnál az alsó könny-meniscusból (a pre-korneális könny film alsó szemhéj okozta horizontális megvastagodása). A mintavétel során figyelmet kellett fordítani arra, hogy a kapilláris a conjunctívát ne érintse. A mintavétel időtartamát mértük és a kalkulált könny szekréción ráta ($\mu\text{l}/\text{min}$) alapján a mintákat két csoportba soroltuk. A vizsgálatokhoz csak azokat a mintákat használtuk fel, ahol a szekréción ráta 5–15 $\mu\text{l}/\text{min}$ értékek között volt. A mintákat a levételt követően azonnal (1800 rpm) 8-10 percig centrifugáltuk. A felülúszót gyorsfagyasztást követően $-80\text{ }^{\circ}\text{C}$ -on tároltuk és kizárólag a kísérletek elvégzéséhez olvasztottuk fel.

3.1.2. Könnyproteomikai analízis

A könnyproteomikai vizsgálatok a Debreceni Egyetem, ÁOK, Biokémiai és Molekuláris Biológiai Intézetében történtek. A módszertan részletes leírása a munkacsoportunk által végzett előkísérletek során született közleményben bemutatásra került [51]. A könny minták vizsgálatához modern nano-HPLC-vel kiegészített ESI-MS/MS tömegspektrométert használtunk. (A vizsgálat részletes bemutatását lásd a hivatkozott közleményben [68].) Munkám során az előkísérletektől eltérően a praeproliferatív és proliferatív retinopathia csoportokat egységes csoportként kezeltem, mivel ebben a kísérletben a cél egy a RD szűrésére szolgáló eljárás vizsgálata volt. A minták globális fehérje koncentráció mintázatát vizsgáltam, ami a könnyben lévő 34 fehérje koncentrációját és azok egymáshoz képesti arányát jelentette.

3.1.3. A gépi tanuló eljárások alkalmazása az automatizált könnyproteomikai alapú szűrőrendszerben

A gépi tanulás a mesterséges intelligencia kutatás egyik területe. A gépi tanuló eljárást alkalmazó informatikai rendszerek tanuló adathalmaz használatával képesek tanulni, majd az új esetek inkomplett adatainak értékelésével megjósolni jövőbeni események bekövetkezését. Jelen esetben az eljárás inputját a könnyfolyadékban mért fehérje koncentrációk értékei jelentik. A tanuló fázisban a rendszer számára megadtuk az esetek kimenetét is, ami jelen esetben a betegek RD státusa (RD/non-RD). A rendszer megtanulta, hogy mely proteomikai koncentráció mintázat jellemzi a RD illetve a non-RD eseteket. Az új eseteknél csak a könnyproteomikai adatok álltak a szoftver rendelkezésére, amely a tanult mintázatok alapján sorolja be az eseteket RD illetve non-RD kategóriába. Mivel az eljárást szűrővizsgálat támogatására fejlesztjük az osztályozó eljárás megvalósítása során elsőrendű kritérium volt a magas szenzitivitási értékek elérése. Ezzel kívánjuk minimalizálni a fals negatív esetek számát csökkentve az esélyét annak, hogy kezelést igénylő esetek ellátás nélkül maradjanak. A projekt során célunk volt különböző osztályozó eljárások teljesítményének értékelése, így ugyanazon az adathalmazon több vizsgálatot is végeztünk különböző osztályozó módszertanokat használva. A modell szelekció során hat különböző modell/algorithmus teljesítményét hasonlítottunk össze Support Vector Machine (SVM), Recursive Partitioning (rpart), Random Forest (randomForest), Naive Bayes (naiveBayes), Logistic Regression (logReg), K-Nearest Neighbor (k-NN).

További célkitűzés volt az egyes markerek (fehérjék) értékelése annak érdekében, hogy kiválaszthassuk azokat, amelyek a leginkább hozzájárultak az osztályozó rendszerünk teljesítményének kialakításához. Az egyes modellek összehasonlíthatóságának érdekében a vizsgálatok során minden osztályozónál azt a paraméterezést használtuk, amelynél a modell teljesítménye a legjobb volt. Az osztályozók a vizsgált adathalmazon elért teljesítményének monitorozásához k-szoros keresztvalidációs eljárást alkalmaztunk [69]. A keresztvalidációs eljárás során az adathalmazt k-számú egyenlő részre osztottuk. Az első k-1 halmazt (tréning adathalmaz) használtuk fel a modellalkotáshoz, melyet aztán a k-adik halmazon tesztelünk (teszt halmaz). Az iterációk során ezt követően újabb k-1 darab részhalmazt alkalmaztunk a rendszert tanítására, majd a maradék k-adik részhalmazon validáltuk a modellt. Az eljárást k-szor hajtottuk végre. A keresztvalidációs ciklusok végén az egyes validációs ciklusok átlagértékét vettük eredményként. A tisztán könnyproteomikai adatokat alkalmazó modell validálása során 5-szörös keresztvalidálást alkalmaztunk. A teljesítmény leírására a standard mérőszámokat használtuk: szenzitivitas, specificitas, pontosság (accuracy), F-érték (a pozitív prediktív érték és a szenzitivitas súlyozott átlaga, mely a teljesítmény mérőszámaként alkalmazható).

Ahogy említettük a gépi tanuló eljárás használatakor a pozitív és "feltehetően pozitív" esetek alkották a RD kategóriát, míg a negatív esetek a non-RD kategóriát. Az eljárás egyfajta "előszűrésenként" lenne alkalmazható a klinikai rutinban. Az előszűrés eredményeként negatívnak ítélt esetek kikerülnek a rendszerből. Ezek az esetek az adott helyen érvényes szűrési protokollnak megfelelően a következő időpontban (többnyire 1 év) ismételt vizsgálatra kerülnek. A pozitív ill. álpozitív esetek digitalizált szemfenéki képei ezzel szemben humán képelemző általi vizsgálatra kerülnek, aki további vizsgálatokra küldi az ellátásra szoruló eseteket, vagyis kiértékesíti, és szemészeti vizsgálatra hívja a betegeket.

3.1.4. Az adatelemzés során alkalmazott szoftver eszközök

A vizsgálat során a feldolgozott könnyproteomikai adatokat és a klinikai adatokat Microsoft Excel szoftverben gyűjtöttük. Ezeket az adatok alkalmaztuk később az elemzések input adataként. Az elemzés során az "R" statisztikai keretrendszer következő programcsomagjait alkalmaztuk: Support Vector Machine: "e1071"; Recursive Partitioning: "rpart"; Random Forest: "randomForest"; Naive Bayes: "naiveBayes"; Logistic Regression: "glmnet" and for visualization purpose: "ggplot2" [70]. A K-Nearest Neighbor modell alkalmazásához munkacsoportunk saját megoldást dolgozott ki. A modellek inputjaként a könnyben mért fehérjék koncentráció értékeit használtuk. Mivel a rendszer felépítésénél elsődleges célnak tekintettük a klinikai rutinban történő alkalmazhatóságot, ezért szerettük volna az input fehérje koncentráció értékek számát csökkenteni. Ennek megfelelően három különböző vizsgálati elrendezést dolgoztunk ki. Ezt követően arra törekedtünk, hogy kiválasszuk az input adatok és a különböző modellek legjobb kombinációját.

Az első vizsgálati elrendezésben 34 fehérje koncentrációját használtuk bemenetként. Ezt követően a második vizsgálati elrendezésben kipróbáltuk, hogy mi történik abban az esetben, amikor a 34 fehérje koncentráció értéke közül csak azon fehérjék koncentráció értékeit használjuk, amelyek az előkísérletek során klasszikus statisztikai módszerek alkalmazásával biomarkerként alkalmazhatónak mutatkoztak szűrőeljárásokban való felhasználásra. Ekkor 6 fehérje koncentrációját adtuk meg a rendszer bemeneteként. Ezt követően a korábbihoz hasonlóan végeztük el a vizsgálatot.

A harmadik vizsgálati elrendezésben PCA-val a 34 fehérjéből kiindulva próbáltuk csökkenteni az input változók számát. Ebben az esetben a dimenzió redukciós eljárást alkalmazva az eredeti adatainkban meglévő információt tömörítettük. Ezzel a módszerrel képesek voltunk az input változók számát csökkenteni úgy, hogy az adatokban meglévő szórás jelentős részben megmaradjon. Így kisebb számú, egymással korrelálatlan bemeneti változóval végeztük el a vizsgálatot, amelyek így az eredeti változók lineáris kombinációjaként szerepeltek az elemzésben.

3.2. A kombinált eljárást alkalmazó szűrés módszertanának vizsgálata során alkalmazott módszerek

3.2.1. Betegvizsgálat és mintavétel

Összesen 52 DM-ban szenvedő beteget vontunk be a vizsgálatok második fázisába (21 férfi; átlagéletkor: 65,2 év; a DM diagnózisától eltelt idő: 16,4 év; 14 NIDDM) a Debreceni Egyetem Szemklinikájának járóbeteg szakrendeléséről. A vizsgálat időpontjában minden egyes beteg antidiabetikus kezelés alatt állt. A betegek közül 39 mutatta a RD jeleit. 16 beteg a 39-ből pedig már egyszer, vagy többször átesett lézeres photocoagulációs eljárásen.

A vizsgálatba bevont betegektől gyűjtött adatokat csak abban az esetben vontuk be a kísérletbe, ahol a vizsgált szemről komplett könnyproteomikai panel és képfeldolgozási módszerekkel elemezhető minőségű digitális szemfenéki fotó is rendelkezésre állt. A potenciálisan vizsgálható 104 szem közül 74 esetben állt rendelkezésünkre minden szükséges könnyproteomikai mérési eredmény és osztályozásra alkalmas fundus fotó. A fennmaradó 30 szem esetén is elvégeztük a klinikai vizsgálatokat, azonban az esetek egy részében a könnyminta vételezés nehézsége tette lehetetlenné a további vizsgálatokat, más részében pedig a retina fotózás volt kivitelezhetetlen. Az előbbi okai leggyakrabban a fentiekben ismertetett kizárási kritériumok voltak, illetve az együtműködés hiánya, összesen 9 esetben. A maradék 21 esetben a fotózás akadályát az üvegtesti vérzés, zárt zugú glaukóma és a cataracta jelentette, ahogyan azt az 1. táblázat is mutatja.

A bevont betegszáma			Vizsgált szemek száma		
52			104		
Non-RD	RD		Vizsgálatba bevont szem	Vizsgálatból kizárt szem	
	Proliferatív	Nonproliferatív		Könnyminta vételezés nehézsége	Retinafotózás nehézsége
13	15	24	74	9	21

1. táblázat. A betegek vizsgálatból való kizárásának okai.

A RD daignózisának megalkotásához standard 7-mezős fundus felvételeket használtunk. A képeket Zeiss készülékkel készítettük: Megaplug Camera Model 1,6i/10 BIT Zeiss (Carl Zeiss Ophthalmic System A6, Jena, Germany). A fundus fotókat ezt követően két független szemész szakorvos vizsgálta. A könnyminták gyűjtése üveg kapillárisal standardizált körülmények között, szakképzett

asszisztens közreműködésével történt a fentebb leírt módon (3.1.1. fejezet). A mintavétel időtartamát rögzítettük. Az elemzéshez csak olyan mintákat használtunk, ahol a szekréciós ráta 5–15 $\mu\text{l}/\text{perc}$ volt.

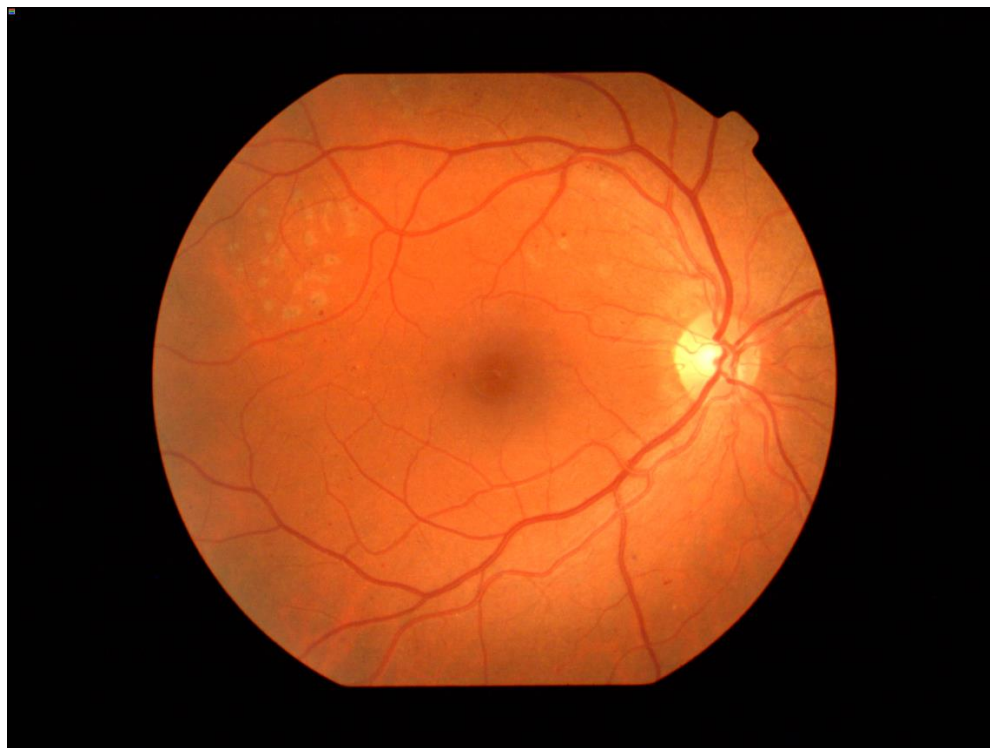
3.2.2. Az alkalmazott képfeldolgozási módszertan bemutatása

A vizsgálatok második fázisában a képfeldolgozáson és könnyproteomikai vizsgálatokon alapuló kombinált eljárást alkalmaztuk. Ennek megfelelően az input adathalmaz a könnyproteomikai kísérletekből származó fehérje koncentráció értékeket, illetve a fundus fotók képfeldolgozási módszertannal történő vizsgálatának eredményeit, a MA-k számát tartalmazta.

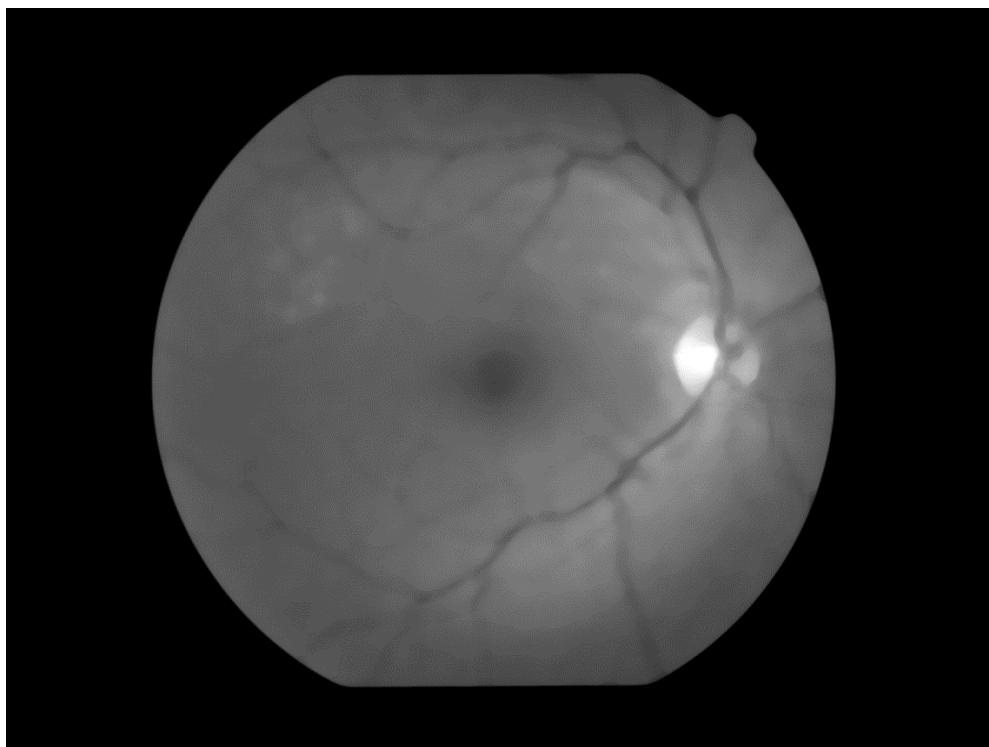
A képfeldolgozási módszertan önálló alkalmazása során első lépésben egy MA detektort hoztuk létre, mely szintén gépi tanulás alapú módszertant alkalmazott a fundusfotókon meglévő MA-k felismerésére. Az ezt követő lépésben – a kombinált eljárásban - az azonosított MA-k száma és a könnyfehérje koncentrációk együttesen jelentették a kombinált módszertan inputját. A kimenetet ez esetben is a RD illetve a non-RD kategóriák jelentették.

MA detektálás

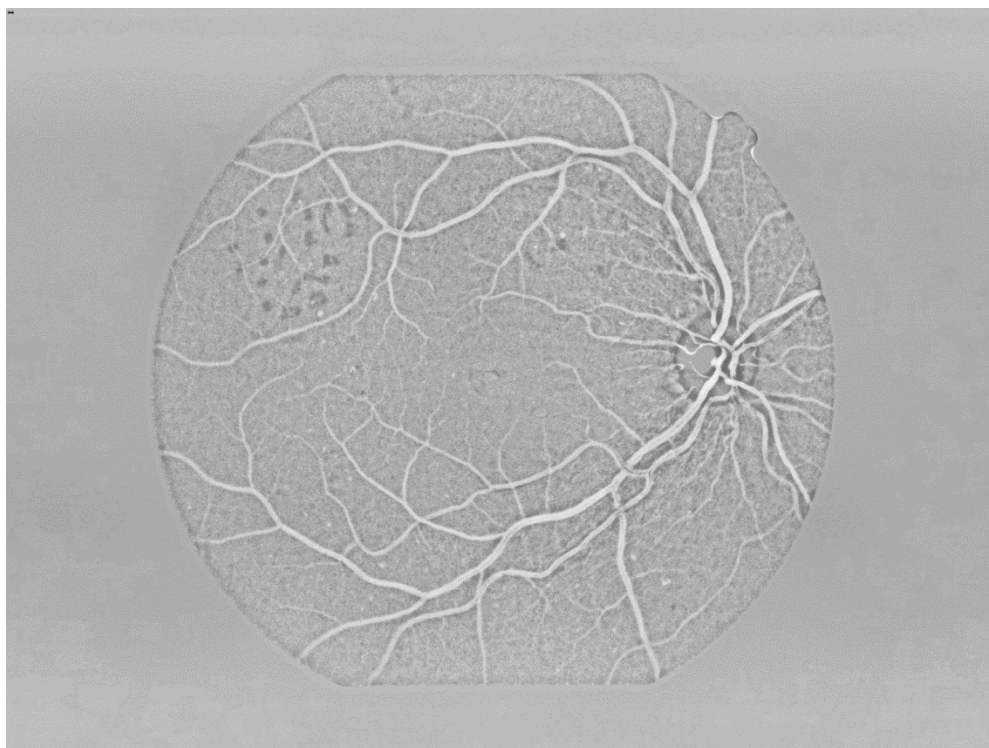
A digitális fundus fotók (13. ábra) képi adata MA detektor segítségével került feldolgozásra. [71,72]. Az alábbi ábrák a képfeldolgozás egyes lépéseit és azok eredményét illusztrálják (14. ábra: medián filterezés; 15. ábra a CLAHE kontraszterősítés).



13. ábra. Az eredeti fundusfotót mutatja feldolgozás előtt.



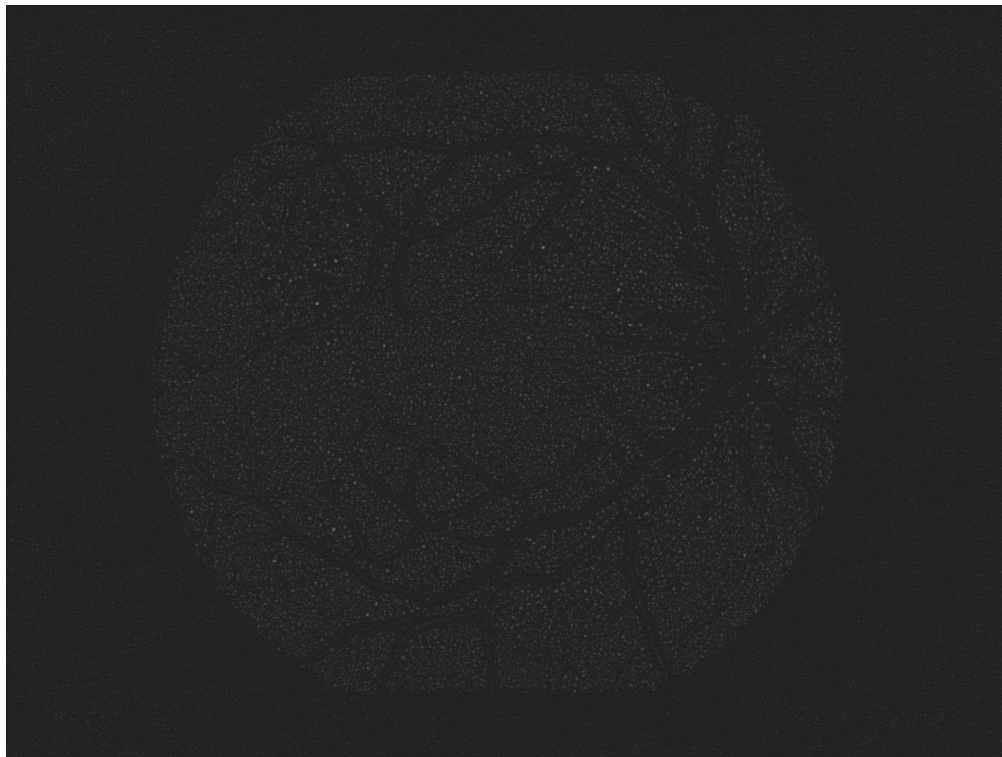
14. ábra. A medián filterezés eredménye.



15. ábra. CLAHE (kontraszt limitált adaptív hisztogram kiegyenlítés) kontraszterősítés.

Mivel a fundus képeken a zöld színcsatorna hordozza a legtöbb információt az érhálózattal kapcsolatban, ezért a háttérkorrekció során ebből került kivonásra a medián filterezett kép (35x35-ös négyzetet használva). 3x3-as medián szűrőt használtunk a képen lévő zajok eltávolítására. A háttér korrigált képen kontraszt-limitált adaptív hisztogram kiegyenlítést (CLAHE) hajtottunk végre, melynek célja a képen lévő képletek és elváltozások kontrasztjainak felerősítése volt.

A képfeldolgozás következő lépése a top hat transzformáció (16. ábra) elvégzése volt morfológiai rekonstrukcióval [73]. A top hat transzformáció egy képfeldolgozási módszer, melyet kisméretű objektumok kivonására használnak. A rekonstruált kép 10x10-es korong-formát használ a kisméretű körkörös elváltozások, vagyis a MA jelöltek azonosítására. A MA jelöltek középpontját kalkuláljuk, majd minden egyes MA jelöltre vonatkozóan megadjuk a következő paramétereket: terület, rotációs inercia, átlagintenzitás a morfológiailag nyitott képen, átlagintenzitás a zöld színcsatornán, szórás a zöld színcsatornán.



16. ábra Top hat transzformáció.



17. ábra. A képfeldolgozást követően láthatóak a MA detektor által azonosított nyers MA jelöltek.

Gradient Boosting Machine (GBM) eljárást használtunk a MA jelöltek osztályozására a fenti jellemzők figyelembevételével [74,75]. A betanításhoz olyan fundus fotókat használtunk, ahol a képen látható MA-k manuálisan be voltak jelölve. A következő lépésben ugyanilyen fotókat használtunk a MA detektor teljesítményének méréséhez.

3.2.3. Könnyproteomikai vizsgálatok

A könnymintákat rögtön a mintavételt követően 8–10 percig centrifugálták (1800 rpm), majd a felülúszó gyorsfagyasztásra került (-80°C). A vizsgálatok a 3.1.2. fejezetben leírtakhoz hasonlóan nano-HPLC-vel kiegészített ESI-MS/MS tömegspektrométerrel történtek [76].

3.2.4. A gépi tanuló eljárás alkalmazása

A projekt második fázisának célja a kombinált, könnyproteomikai és képi adatokat is alkalmazó módszertan vizsgálata volt. A vizsgálat során három modellt építettünk fel.

- Az első modellben a képfeldolgozáson alapuló módszertan egyéni teljesítményét vizsgáltuk.

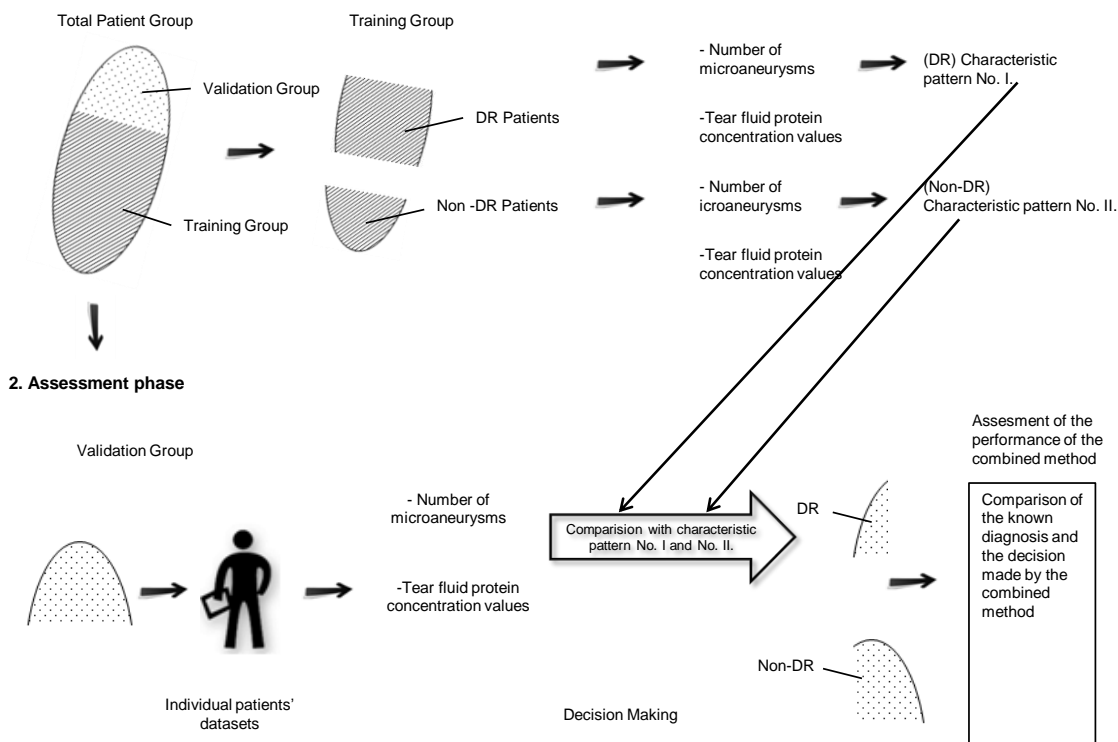
- A második modellben a kizárólag könnyproteomikai adatok használatára épülő modellt értékeltük.
- A feltételezésünk az volt, hogy a két módszertan kombinálásával létrehozott harmadik modell teljesítménye meghaladja az első két modellt.

A projekt során minden esetben a legjobban teljesítő osztályozó eljárást alkalmaztuk. A képfeldolgozáson alapuló modell esetén naive Bayes osztályozót, a könnyproteomikai adatokat használó illetve a kombinált modell esetén pedig GBM-típusú osztályozó eljárást. A teljesítmény méréséhez ez esetben 10-szeres keresztvalidációs eljárást (10-szer ismételve) használtuk a fentiekben (3.1.3. fejezet) leírt módon.

A fundus képeket a bemutatott MA detektorral dolgoztuk fel, melynek során egy küszöbértéket állítottunk be, amit a potenciális MA-k osztályozásának céljára használtunk. Az első modell esetén a MA detektor által azonosított MA-k száma volt az egyetlen jellemző, mely alapján a gépi tanuló eljárás RD, vagy non-RD csoportba sorolta az adott képet (18. ábra). Az osztályozáshoz naive Bayes osztályozót alkalmaztunk. Az elemzéseket és az osztályozást nem a beteg, hanem a vizsgált fénykép szintjén végeztük. A második modellben 34 könnyfehérje koncentrációjának globális mintázatát és annak változását vizsgáltuk. A kombinált modellben a fundus képeken a MA detektor által azonosított MA-k száma és az ugyanazon szemhez tartozó könnymintában mért fehérjék koncentrációja volt a rendszer bemenete.

A tanulás fázisában emellett megadtuk az adott szem RD státuszát is. A tanulási fázist követően a tesztelési fázisban az azonosított MA-k számát és a fehérje koncentrációs értékeket adtuk meg a szem RD státusza nélkül. Az osztályozó feladata pedig az volt, hogy a rendelkezésére álló inkomplett információk alapján próbálja besorolni az eseteket a megfelelő kategóriába. Célunk annak igazolása volt, hogy a kombinált eljárás teljesítménye meghaladja a másik két modell egyéni teljesítményét. A 18. ábrán a gépi tanuló eljárás alkalmazásának tanuló és értékelési lépései láthatók a kombinált módszertan példáján.

1. Learning Phase



18. ábra. A gépi tanuló eljárás alkalmazása a kombinált módszertan példáján bemutatva. Tanuló fázis (fent): Az esetek (szem) halmazából véletlenszerűen választottunk ki egy szubpopulációt, amit tanuló halmaznak nevezünk, majd az ismert diagnózisok alapján a szubpopuláció eseteit két csoportra osztottuk RD-re ill. non-RD-re. A klinikai diagnózis (RD/non-RD), a retina képeken azonosított MA-k száma, ill. a fehérje koncentráció értékek jelentették a gépi tanuló algoritmus inputját. Az algoritmus képes megtanulni, hogy mely mintázatok a leginkább jellemzőek a RD ill. a non-RD csoportra. Tesztelési fázis (lent): A következő lépésben a másik szubpopuláció (maradék esetek), a teszt halmaz adatait használtuk. Ebben a lépésben csak a fehérje koncentrációk értékeit illetve a MA-k számát tekintettük az algoritmus inputjaként. A szem RD státuszát ebben a lépésben nem adtuk meg bemenetként. A tanuló algoritmus összevetette az új input jellemző mintázatait azokkal a RD ill. non-RD esetekre jellemző mintázatokkal, amelyeket a tanuló fázisban megtanult. Ez alapján az algoritmus döntést hoz a teszt halmaz eseteire vonatkozóan. A modell teljesítményének értékeléséhez ezt az output-ot összevetjük az ismert klinikai diagnózissal.

A módszer értékeléséhez k-szoros keresztvalidációs eljárást alkalmaztunk a fentebb leírtaknak megfelelően, azonban a k ebben az esetben 10 volt. Az értékelési ciklusokat egymás után 10-szer ismételtük. Teljesítmény indikátorként ez esetben is a specificitás, szenzitivitás, pontosság és az F-érték szolgált. A 10-szeres keresztvalidáció során a tanuló és a teszt halmaz eseteit random módon soroltuk be egyik, vagy másik csoportba.

Bizonyos esetekben egy beteg egyik szeme az egyik a másik szeme a másik halmazba (tanuló/teszt) került besorolásra. Figyelembe véve a vizsgálatba bevont személyek relatíve alacsony számát, ez a besorolási módszer szignifikánsan javíthatna volna a modellek teljesítményét. A probléma elkerülésének érdekében ezért ilyen esetekben a szempár tanuló halmazba eső elemét kizártuk a vizsgálatból.

3.2.5. Az adatelemzés céljára szolgáló szoftver eszközök

A klinikai és kísérleti (proteomikai) adatok gyűjtéséhez a SmartBiobank adatmenedzsment rendszert használtuk (AstridBio Technologies Kft.) [77]. Az adatelemzés céljára az R statisztikai programcsomag következő csomagjait használtuk: „gbm”, „caret”, „stat”, „e1071” [78]. A képfeldolgozási feladatokat az Octave szoftverben végeztük el [79].

A jelen munkában bemutatásra kerülő proteomikai, illetve kombinált proteomikai és képfeldolgozási módszertanon alapuló eljárásokhoz hasonló módszer kidolgozásával ez idáig más kutatócsoport nem próbálkozott.

3.3. Etikai engedély, betegbeleegyezés

A jelen munkában leírásra került vizsgálatok során figyelembe vettük a ránk vonatkozó, érvényes nemzetközi és hazai jogi és etikai szabályozást. A betegvizsgálatok a bevont személyek tájékozott beleegyezésével kerültek elvégzésre. A vizsgálatok etikai engedélyének száma: IX-R-052/00016-26/2012. A vizsgálat összhangban állt a Helsinki deklarációban (1964) leírtakkal.

3.4. Archiválás

A nyers és köztes adatokat Excel táblázatokban, a szemfenéki képekkel együtt duplikálva, hordozható és beépített merevlemezzen tároltuk.

4. Eredmények

4.1 A könnyproteomikai alapú eljárás értékelésével kapott eredmények.

4.1.1. Hat különböző gépi tanuló eljárás összehasonlítása

A hat különböző gépi tanuló eljárás közül az rpart bizonyult a leghatékonyabbnak, ha kizárólag a klasszikus statisztikai eljárásokkal azonosított hat marker fehérje adatait használtuk. Amint az a 2. táblázatban látható, az értékelési mérőszámok azt mutatják, hogy a recursive partitioning (rpart)

osztályozó használatán alapuló modell teljesítménye felülmúlja a másik öt modell teljesítményét (0,74 szenzitivitás, 0,48 specificitás és 0,65 pontosság; rpart/marker). Ugyan a Naiv Bayes osztályozó érte el a legmagasabb szenzitivitás értéket (0,80; naiveBayes/marker), de a specificitása rendkívül alacsony (0,38). A Random Forest osztályozón alapuló modell kissé rosszabbul teljesített, mint a rpart.

Modell	Dataset	SENS	SPC	ACC	PREC	NPV	F1	LRP	LRN
naiveBayes	orig	0.6991	0.4186	0.6218	0.7596	0.3462	0.7281	1.2025	0.7188
	marker	0.8000	0.3874	0.5064	0.3462	0.8269	0.4832	1.3059	0.5163
	pca	0.6731	0.3365	0.4487	0.3365	0.6731	0.4487	1.0145	0.9714
kNN	orig	0.6711	0.5000	0.6667	0.9808	0.0385	0.7969	1.3421	0.6579
	marker	0.6688	0.5000	0.6667	0.9904	0.0192	0.7984	1.3377	0.6623
	pca	0.6643	0.3077	0.6346	0.9135	0.0769	0.7692	0.9596	1.0909
logReg	orig	0.6923	0.3846	0.5897	0.6923	0.3846	0.6923	1.1250	0.8000
	marker	0.6615	0.3077	0.6026	0.8269	0.1538	0.7350	0.9556	1.1000
	pca	0.6623	0.0000	0.6538	0.9808	0.0000	0.7907	0.6623	Inf
randomForest	orig	0.6929	0.4483	0.6474	0.8462	0.2500	0.7619	1.2559	0.6850
	marker	0.6923	0.4103	0.6218	0.7788	0.3077	0.7330	1.1739	0.7500
	pca	0.6748	0.3636	0.6090	0.7981	0.2308	0.7313	1.0604	0.8943
rpart	orig	0.7083	0.4722	0.6538	0.8173	0.3269	0.7589	1.3421	0.6176
	marker	0.7404	0.4808	0.6538	0.7404	0.4808	0.7404	1.4259	0.5400
	pca	0.6935	0.4375	0.6410	0.8269	0.2692	0.7544	1.2330	0.7005
SVM	orig	0.6645	0.0000	0.6603	0.9904	0.0000	0.7954	0.6645	Inf
	marker	0.6623	0.0000	0.6538	0.9808	0.0000	0.7907	0.6623	Inf

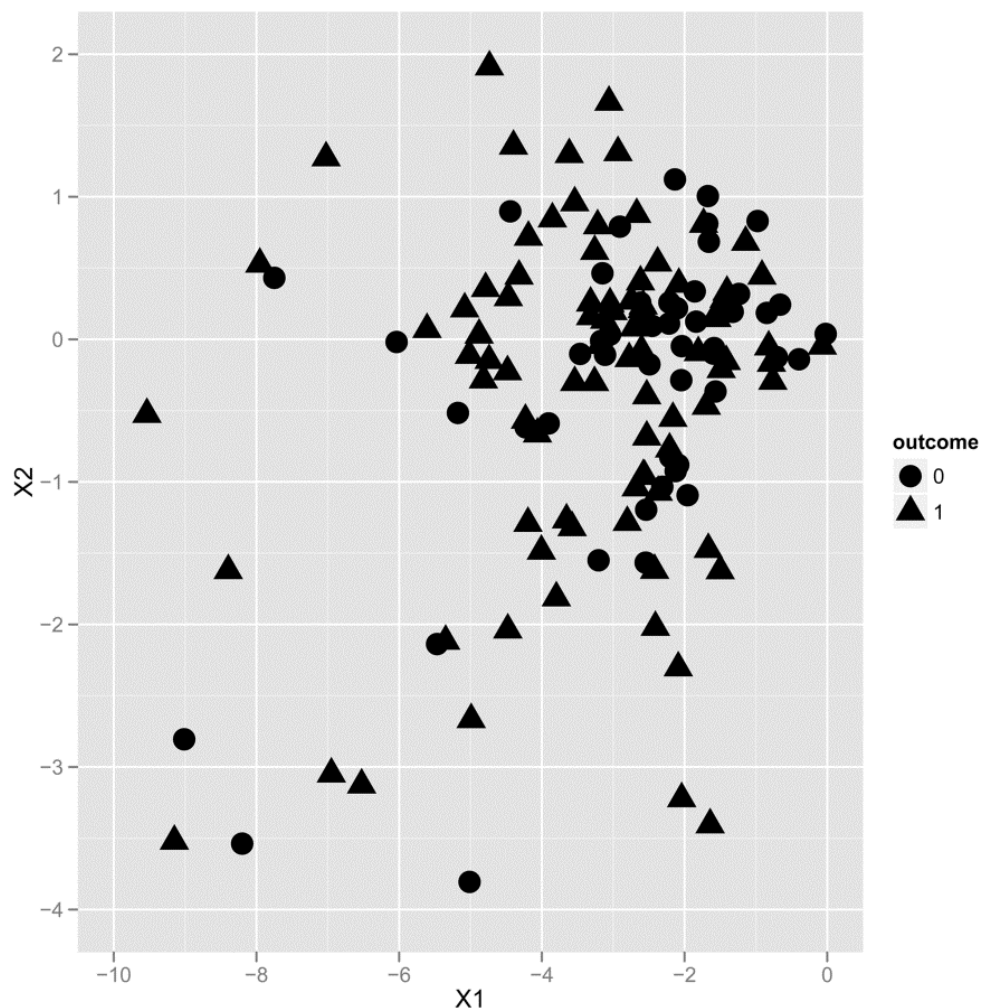
pca	0.6623	0.0000	0.6538	0.9808	0.0000	0.7907	0.6623	Inf
-----	--------	--------	--------	--------	--------	--------	--------	-----

2. táblázat. A hat különböző osztályozó teljesítmény mutatóinak összevetése. A hat osztályozó értékelése során használt teljesítményindikátorok a különböző inputok elemzése során (globális proteomikai mintázat); a 6 marker fehérje; és a főkomponens analízis (PCA) alkalmazása esetén: orig - teljes adathalmaz; marker - csak a marker fehérjék; pca – PCA által transzformált adatok. Az oszlopok megnevezései: SENS – szenzitivitás; SPC – specificitás; ACC – pontosság; PREC - pozitív prediktív érték; NPV - negatív prediktív érték; F1 - F-érték; LRP – pozitív valószínűségi hányados; LRN – negatív valószínűségi hányados.

A projekt kezdetén a hat fehérjét független biomarkerként azonosítottuk a 34 vizsgált fehérje közül statisztikai hipotézis vizsgálat segítségével.

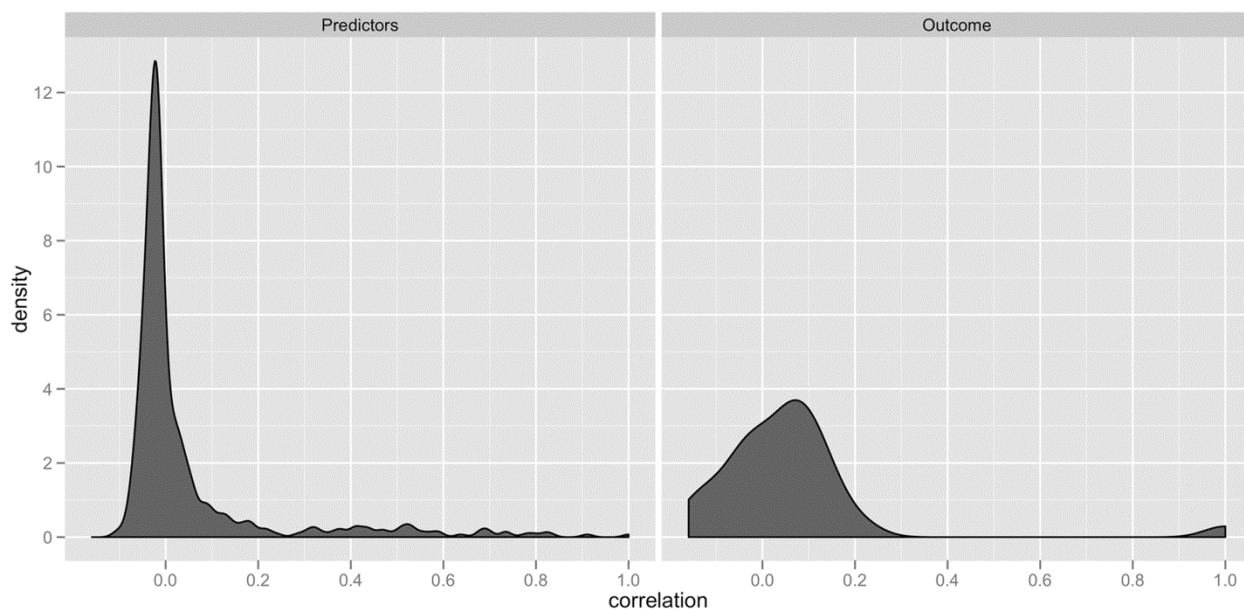
A globális mintázat figyelembevételével történő vizsgálódás motivációját az adta, hogy feltételeztük, a markerként nem azonosított fehérjék is hordoznak olyan információt, amelynek felhasználása javíthatja a csak a hat marker használatával elérhető teljesítményt. A PCA alkalmazásának, vagyis a tanításhoz felhasznált adatmennyiség (input) csökkentése mögött az a törekvés állt, hogy egy a rutinban is alkalmazható, a teljes proteomikai panel tömegspektrométerrel való mérésénél egyszerűbben és olcsóbban megvalósítható eljárást dolgozzunk ki. Emellett szerettük volna megőrizni az eredeti eljárás teljesítmény értékeit. A vizsgálat eredményeként azt állapítottuk meg, hogy sem a 6 marker fehérjén, sem a PCA alkalmazásával elvégzett vizsgálat esetén nem kaptunk lényegesen eltérő eredményeket, mint amit a teljes proteomikai panel vizsgálatával kaptunk. A PCA során kapott első két komponens hordozta a teljes fehérje panel koncentráció értékei varianciájának 22%-át.

A 2. táblázatban látható, hogy ha kizárólag az első két komponens tartjuk meg a modell teljesítménye nem változik jelentősen. Az adatok vizualizálásához kigeneráltuk az első két komponens scatter plot ábráját. A 19. ábrán látható, hogy a két osztály nem különül el élesen egymástól.



19. ábra. A PCA transzformált adathalmaz Scatter plot-ja. X1 és X2 a transzformáció után megtartott két komponens. A 0 a non-RD, az 1 a RD állapotra utal.

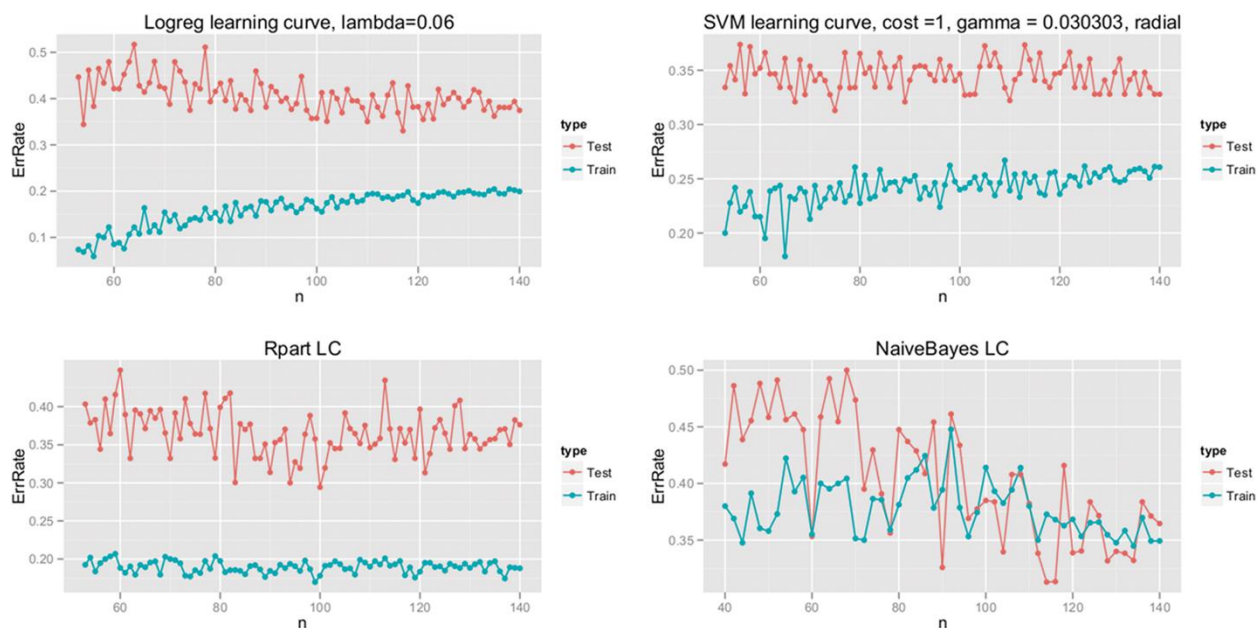
A leíró elemzésünk részeként, hogy pontosabb képet kapjunk az adathalmazról, értékeltük és vizualizáltuk a prediktor változók (fehérje koncentrációs értékek), valamint a prediktor változók és a kimeneti változók (RD/non-RD) közötti összefüggést. A korrelációs értékek sűrűségfüggvényén (20. ábra) látható, hogy a prediktor és a prediktor/kimenet között nagyon alacsony korreláció mutatkozik.



20. ábra. A valószínűségi változó sűrűségfüggvénye. A prediktorok közötti korrelációs értékek sűrűségfüggvénye látható a kép bal oldalán. A jobb oldalon a prediktor és a kimeneti változók közötti korrelációs értékek sűrűségfüggvénye látható.

4.1.2. Az adathalmaz méretének hatása a modell teljesítményére

Annak érdekében, hogy vizsgálni tudjuk a tanuló adathalmaz méretének hatását az eredmények megbízhatóságára, összehasonlítottuk minden egyes eset tanulási görbáját. Ezzel a módszerrel rálátást kaptunk az egyes modellekre jellemző torzításra és szórására, valamint ötleteket kaphattunk a modell jövőbeni javítási lehetőségeire vonatkozóan. A teljesítmény pontosabb értékeléséhez meg kellett vizsgálnunk, hogy van-e hatása az adathalmaz növelésének a különböző gépi tanuló eljárások teljesítményére (21. ábra).



21. ábra. Tanulási görbék. A Support Vector Machine (SVM), Recursive Partitioning (Rpart) Logistic Regression (Logreg) és a Naive Bayes (naiveBayes) osztályozók tanulási görbéje különböző méretű adathalmazokra vonatkozóan.

Tanulási görbék ábrázolásával szemléltettük a tanulási és a tesztelési lépés során tapasztalható hibát, különböző méretű tanulóhalmazok mellett. A görbék egymástól való távolsága mutatja a különbséget a tanuló és a teszt halmaz esetén tapasztalt hibaszázalékok között, amely a Support Vector Machine (SVM) és a Recursive partitioning (Rpart) algoritmusok esetében nagy. Az eredmények azt mutatják, hogy ezen osztályozók esetén a szórása nagy, ezért a tanuló adathalmaz méretének növelése hozzájárulhat a teljesítmény növeléséhez. A Logistic regression (Logreg) és a Naive Bayes (naiveBayes) osztályozók esetén viszont nem látunk számottevő különbséget a tanulási görbék között, azonban a tanuló és teszt halmaz esetén is magas a hibaszázalék, ami azt jelentheti, hogy a modellünk túlságosan egyszerű.

4.2. A kombinált módszertan alkalmazásával kapott eredmények

A kombinált módszer alkalmazásával azt kívántuk bizonyítani, hogy két különböző típusú adatforrásból érkező adatfésülés kombinált használatával az általunk felépített RD szűrésre szolgáló módszertan jobb teljesítményre képes, mint amire az egyes adatfésülések felhasználásával felépített modellek külön-külön képesek. A képfeldolgozási módszertanon alapuló szűrőrendszereknek jelentős szakirodalma van, amelyből tudható, hogy az ezeken a módszertanokon alapuló rendszerek teljesítménye

a humán képelemzők teljesítményét közelíti. Ezt a teljesítményt kívántuk tovább javítani a proteomikai és a képfeldolgozási adatok kombinálásával.

Ennek megfelelően három modell teljesítményét hasonlítottuk össze. Az első modellben a MA-k száma volt az egyetlen jellemző, amelyet az osztályozó algoritmus értékelt, mielőtt a RD illetve a non-RD csoportba sorolta a vizsgált fundus képeket. Az MA detektor önállóan 0,84-es szenzitivitás illetve 0,81-es specificitás értéket ért el. A második modell, amely kizárólag proteomikai adatokra támaszkodott 0,87-es szenzitivitás értéket ért el 0,68-as specificitás mellett. A kombinált eljárás teljesítménye meghaladva mindkét vizsgált eljárás teljesítményét 0,93-as szenzitivitás és 0,78-as specificitás értéket ért el. Ez feltételezhetően annak köszönhető, hogy a két eltérő adatfésülés egymástól független, de kiegészítő információt biztosított a kimenetekről az osztályozáshoz.

A 3. táblázatban bemutatott adatokon látható, hogy az MA detektor szignifikánsan jobban teljesített, mint a proteomikai alapú módszertan. A pontosság tekintetében a kombinált módszertan az első és a második modellnél is jobban teljesített. A MA detektor specificitási mutatói jobbak, mint a kombinált módszertané, azonban - mivel a munka célja egy olyan elő-szűrő eljárás kifejlesztése volt, melyet egy humán képelemzők általi szűrés követ a pozitívként (illetve álpozitívként) kiszűrt betegek esetében - a szenzitivitás növelése az elsődleges célunk még akkor is, ha ez kompromisszummal jár a specificitás tekintetében. Az alacsonyabb specificitás a gyakorlatban azt jelenti, hogy nagyobb betegpopulációt kell az elő-szűrést követően a humán képelemzőknek megvizsgálniuk.

Szűrő eljárás		SENS	SPC	ACC	PREC	NPV	F1	LRP	LRN
Képfeldolgozás	Átlag	0.84	0.81	0.84	0.94	0.63	0.89	4.42	0.20
	SD	0.11	0.04	0.10	0.13	0.11	0.11	2.36	0.11
Proteomika	Átlag	0.87	0.68	0.82	0.89	0.63	0.88	2.72	0.19
	SD	0.17	0.12	0.11	0.21	0.15	0.16	1.24	0.12
Kombinált eljárás	Átlag	0.93	0.78	0.89	0.93	0.78	0.93	4.23	0.09
	SD	0.18	0.19	0.15	0.23	0.22	0.18	1.32	0.07

3. táblázat. A szűrőeljárások teljesítmény mutatóinak bemutatása. A képfeldolgozás, a proteomikai illetve a kombinált szűrőeljárás teljesítmény mutatóinak bemutatása. Az oszlopok jelentése: SENS – szenzitivitás; SPC – specificitás; ACC - pontosság; PREC - pozitív prediktív érték; NPV - negatív prediktív érték; F1 - F-érték; LRP – pozitív valószínűségi hányados; LRN – negatív valószínűségi hányados.

4.3. Új eredmények összefoglalása

4.3.1. 1.munkafázis:

- (1.) A különböző könny fehérjék koncentrációinak globális mintázatát gépi tanulás alapú módszerekkel elemezve, sikerrel azonosítottunk olyan mintázatokat, amelyek a RD-ban szenvedő betegekre jellemzőek voltak.
- (2.) Olyan eljárást alkottunk, amely a könnyfehérje koncentrációk értékeit figyelembe véve, (kísérleti modellben) a DM-ban szenvedő betegeket DR és non-DR kategóriákba volt képes osztályozni.
- (3.) Sikerrel hoztunk létre olyan szűrőeljárást, amely az előkísérletek során a kutatócsoport által klasszikus statisztikai módszerekkel azonosított 6 markerfehérje koncentrációinak egyidejű használatával képes a szűrt populáció DM-ban szenvedő betegeinek osztályozására (RD, illetve non-RD csoportokba).

4.3.2. 2. munkafázis:

- (4.) A területen elsőként sikerrel kombináltunk képfeldolgozás és könny fehérje proteomika alapú eljárást a DR szűrésére (kísérleti modellen).
- (5.) A kombinált eljárás teljesítménye meghaladta az egyedi (képfeldolgozás, ill. könnyproteomikai eljárás) eljárások teljesítmény mutatóit, vagyis a független, eltérő forrásból származó információk együttes használata egy az egyedi eljárásoknál erősebb osztályozót eredményezett.

5. Megbeszélés

A munka célkitűzése egy olyan eljárás kidolgozása volt, amely alkalmas a diabeteses betegek, mint a RD kockázatának kitett populáció, elő-szűrésének végrehajtására. A negatív esetek kiválogatásával csökken az a betegszám, amelyet a szűrőcentrumokban dolgozó képelemzőknek kell megvizsgálnia. Mivel a jelenlegi szűrőeljárások igen munkaerő-igényesek és így költségesek, a szűrni kívánt populáció csökkentése jelentős költségmegtakarítással járhat a jövőben. Az elmúlt időszakban számos olyan vizsgálatot végeztek és publikáltak különböző munkacsoportok, amelyben képfeldolgozó eljárásokat dolgoztak ki RD szűrésben való alkalmazás céljával. Ezek a módszerek ígéretes teljesítményük ellenére csak igen lassan terjednek a klinikai gyakorlatban. Ezekkel a közleményekkel párhuzamosan közölte saját munkacsoportunk a könny folyadék proteomikai analízisén alapuló szűrő eljárását.

Az önállóan alkalmazott könnyproteomikai alapú szűrés, mivel a könnyminta vétel nem invazív és viszonylag könnyedén végrehajtható, a háziorvos rendelőjében, vagy akár a beteg otthonában is elvégezhető. A vizsgált személy otthonában, a vizsgált személy által kivitelezett mintavételre láthatunk példát napjainkban. Ilyen a direct to customer (DTC) genetikai tanácsadás keretében végrehajtott nyálminta vétel [80]. Esetünkben a beteg a saját otthonában, maga által levett könnymintát postán

elküldheti a szűrőközpont laboratóriumába, így kizárólag a kiszűrt pozitív, illetve ál-pozitív eseteknek kell a szűrőcentrumba látogatniuk.

A vizsgálatok során a modell által elért teljesítmény értékek alapján úgy ítéljük meg, hogy a módszer jelenlegi formájában még nem alkalmas előszűrésként való alkalmazásra a klinikai rutinban. Mivel ez a vizsgálat az első a maga nemében, még számos lehetőség van a módszer javítására. További vizsgálatokban, mint azt az "eredmény" fejezetben láttuk, elengedhetetlen a tanuló adathalmaz méretének növelése, melynek köszönhetően a rendszer sokkal precízebben lesz képes megjósolni a vizsgálat kimenetelét. Az általunk végzett vizsgálatok során az alkalmazott MS/MS módszertan költségei egyelőre akadályozták az olyan, nagyobb betegszámon történő vizsgálatokat, mint amelyek a képfeldolgozás alapú vizsgálatoknál megszokottak.

A nagy-áteresztőképességű technológiák költségeinek rohamos csökkenése hamarosan elérhetővé teszi alkalmazásukat a napi klinikai rutin számára is [81]. Ezzel párhuzamosan a képzett munkaerő költsége fokozatosan növekszik, amellett, hogy számos helyen a megfelelően képzett munkaerő nem is áll rendelkezésre. Ezen tényezők együttesen a nagy-áteresztőképességű technológiák gyors elterjedését valószínűsítik a klinikai gyakorlatban.

A RD szűrés tekintetében, habár számtalan jól működő *képfeldolgozás alapú módszert* publikáltak ezidáig, melyek teljesítménye a humán vizsgálatnál összemérhető teljesítményű, nagyon alacsony azoknak a rendszereknek a száma, amelyeket független kohorszon validáltak, nemzetközileg elfogadott standardokat követve. Az Abramoff és munkatársai által kidolgozott módszertan többféle képfeldolgozás alapú módszertant kombinál. A rendszerük MA detektor mellett a bevérvések, a gyapottépésszerű gócok, és az kemény exudatumok jelenlétét detektálja egy 874 DM-ban szenvedő beteget tartalmazó populáción. A módszertant irreguláris laesiok, mint a nagy bevérvések illetve a neovascularisatio, felismerésére is felkészítették. A módszer szenzitivitása 96,8%-osnak (95% CI, 94,4%–99,3%) adódott 59,4%-os (95% CI, 55,7%–63,0%) specificitás mellett. A relatíve alacsony specificitás értékek a rendszer használhatóságát jelentősen korlátozzák [82]. A Gotman és munkatársai által végzett vizsgálat célkitűzése a biztosan negatív, normál képek eltávolítása a képatadbázisból a humán vizsgálok által végzett szűrést megelőzően csökkentve ezáltal a vizsgálandó képatadbázis méretét, így a szűrés humánerőforrás igényét. A kombinált képfeldolgozás alapú módszer MA detektálás mellett blot haemorrhagiák (BH) és exudatumok (EX) azonosítására lett kidolgozva. Céljuk a 100%-os szenzitivitás elérése volt, amely mint ahogy az előre látható volt, alacsony specificitás értékkel párosul. A módszertan elő-szűrő eljárásként alkalmazva a képek 26,4% (MA/BH/EX, két mezős vizsgálat), 38,1%-át (MA, macula centrált kép vizsgálata) volt képes eltávolítani a szűrésre kerülő képatadbázisból (mint biztosan

negatív esetet). Ez a munkaterhelés 1/3-os csökkenését jelenti [83]. A saját eljárásunk 100%-os szenzitivitás-beállítás melletti vizsgálata számunkra is érdekes eredményeket hozó új irány lehet.

A munkacsoportunk által létrehozott modellek alkalmasak arra, hogy a proteomikai adatokkal párhuzamosan egyéb adatfészeségeket is inputként használjunk a vizsgálatok során. Ilyenek lehetnek például az anamnézisben szereplő klinikai adatok, vagy képi adatok, mint ahogy azt a saját munkánk második fázisának leírásánál láthattuk. Ha emeljük az inputként felhasznált adatfészeségek számát az szignifikánsan emelheti az eljárás teljesítményét.

Annak ellenére, hogy a módszer jelenlegi formájában nem alkalmas a klinikai gyakorlatban való alkalmazásra, a módszer egyértelműen alkalmas lehet a képfeldolgozás alapú eljárások teljesítményének javítására [84]. A különböző adatfészeség együttes használatának kedvező hatását a predikció pontosságára saját tapasztalatunk mellett a szakirodalom is megerősíti [85,86]. A megnövelt teljesítmény az egyik faktor, ami hozzájárulhat a módszer elterjedéséhez. A másik tényező a mintavétel könnyű kivitelezhetősége, ami lehetővé teszi a tömeges RD szűrés megszervezését vidéki, ritkán lakott, vagy rossz infrastruktúrával rendelkező, vagy nehezen elérhető területeken. Kombinált diagnosztikai módszertanokat más területeken már alkalmaznak különböző szűrő eljárások teljesítményének javítására. Ilyen például az MR felvételek és a genetikai tesztek (BRCA1 mutáció) kombinált alkalmazása az emlőtumor szűrése során [87]. A képfeldolgozás és a könnyű fehérje alapú markerek két jelentősen eltérő adatforrást reprezentálnak így a RD-es szemre nézve eltérő típusú információt hordoznak, aminek köszönhetően egymás hibáinak javítására képesek.

A kombinált eljárással általunk, vizsgálati beállítások között elért 0,93-as szenzitivitás és 0,78-as specificitás közelíti a humán vizsgáló által végzett szűrés megfelelő értékeit. Ahogy fentebb említettük, jelenleg a RD szűrés által elérendő célértékek a 0,80-as szenzitivitás és 0,95-as specificitás. Figyelembe véve, hogy a vizsgált szemek közül 9-et a könnyproteomikai mintavétel problémái, 21-et pedig a fotózás nehézségei miatt ki kellett zárunk a további vizsgálatból, a klinikai vizsgálati protokoll javításán mindenképpen dolgoznunk kell a jövőben.

A jövőre nézve további célunk a proteomikai kísérletben vizsgált fehérjék számának csökkentése. Ha ezt a célt sikerrel valósítjuk meg, a MS/MS módszertant egyszerű fehérje assay-re cserélve a szűrővizsgálatot jelentősen egyszerűbbé, gyorsabbá és költséghatékonyabbá tehetjük. A PCA módszere jól használható az ideális, legjobb predikciós értékkel bíró marker fehérjék kiválasztására. A kiválasztott fehérjék vizsgálata során szeretnénk a koncentráció meghatározás helyett egy küszöbértéket megadni, amelynek köszönhetően a fehérje assay-nek csak egy igen/nem választ kell a rendelkezésünkre bocsátania a szűrés megvalósítása során. A kifejlesztett gyorsteszt a képfeldolgozási módszerrel egy időben is

alkalmazható eljárás lehet. A két vizsgálat kombinálása az egyedi vizsgálatoknál jobb eredményt adhat, ami hozzájárulhat a módszer gyakorlati adaptálásához.

Mivel a módszertanban rejlő lehetőségek vizsgálatának még csak a kezdetén járunk a módszer teljesítménye minden bizonnyal tovább javítható a jövőben. Az eredmények javítására számos elképzelésünk van, melyek közül az alábbi három tűnik a legígéretesebbnek: i) mind a könnyproteomikai, mind a képfeldolgozás alapú módszer javítható a gépi tanuló eljárás paraméterezésének finomhangolásával. Állítható például a MA detektálásához beállított küszöbérték, vagy a vizsgált fehérjék száma stb.; ii) egy következő vizsgálatban szelektálhatjuk a kombinált módszertan számára legideálisabb osztályozó algoritmust; iii) feltételezzük, hogy a betegvizsgálati és a labordiagnosztikai protokollok is tovább javíthatóak.

Figyelembe véve, hogy a terület szakembereinek célja jelenleg nem a humán vizsgáló kiváltása, hanem csak munkaterhének csökkentése, a jövőben hangsúlyt fektethetünk a szenzitivitás 100%-os beállítása mellett a specificitás növelésére. Ezzel számíthatóvá válik a munkateher csökkentésével előidézett költségcsökkenés, ami alapján kalkulálható, hogy az általunk végzett vizsgálat költsége hogyan viszonyul a megtakarítások nagyságához, vagyis a vizsgálat költséghatékonysága meghaladja-e a jelenleg alkalmazott eljárásét (humán vizsgálat). Mivel várhatóan az ember által végzett vizsgálat költségei nőni fognak, a laboratóriumi illetve az informatikai technológiák pedig egyre olcsóbbá válnak, az idő egyértelműen a miénkhez hasonló eljárások kifejlesztésének kedvez. A mi módszerünk tovább fejlesztett változata, vagy egy ezzel párhuzamosan kifejlesztett alternatív technológia minden bizonnyal hamarosan megjelenik majd a RD szűrés napi gyakorlatában.

6. Összefoglalás

6.1. Magyar nyelvű összefoglaló

Könnyproteomikai és képfeldolgozási technikák alkalmazhatóságának vizsgálata a retinopathia diabetica szűrésében.

A RD a cukorbetegség egyik leggyakoribb szövődménye, mely az összes vakságban szenvedő személyek 5%-ánál tehető felelőssé az állapot létrejöttéért. Azoknál a 2-es típusú DM-ban szenvedő betegeknél, akiknél a betegség 10 éve fennáll, 80% az esélye, hogy valamilyen fokú RD már kialakult. Mivel a RD kezdetben tünetmentes, a diagnózis időpontjában a betegség sokszor már igen előrehaladott stádiumban van. Időben történő diagnózissal és terápiával a RD progressziója jelentősen lassítható, így a DM-ban szenvedő populáció számára a rendszeres szemészeti szűrővizsgálaton (legalább évente) való

részvétel ajánlott. A RD szűrés humánerőforrás igénye, így költségei is igen magasak, ami a szűrés automatizálásával csökkenthető lenne. Jelen munka célja a RD szűrésére alkalmas automatizált eljárás létrehozása.

Az első munkafázisban olyan módszertant dolgoztunk ki, amely a könny fehérjék RD okozta koncentráció változásainak vizsgálatán alapszik. A DM-os betegek osztályozásához (RD/non-RD) gépi tanuló eljárást használtunk, mely a fehérje koncentrációk globális mintázatának megváltozása alapján sorolta a betegeket egyik, vagy másik csoportba. A vizsgálatok eredményeként kapott viszonylag alacsony szenzitivitás (0,74) és specificitás (0,48) értékek alapján arra a következtetésre jutottunk, hogy a módszer jelenlegi állapotában önállóan még nem érett a klinikai rutinban való alkalmazásra. Azonban más típusú eljárással kombináltnak használva alkalmas lehet a másik módszer teljesítményének javítására.

A második munkafázisban a könny proteomika alapú módszertant digitalizált retina képek automatizált feldolgozásával kombináltuk. A vizsgálatok során a képfeldolgozás alapú módszer önállóan 0,84-es szenzitivitás és 0,81-es specificitást ért el. A könnyproteomikai módszertan ebben az esetben 0,87-es szenzitivitás és 0,68-as specificitást produkált. A kombinált eljárás inputját a digitális fundus fotón azonosított MA-k száma és a könnyfehérje koncentrációs értékek jelentették. Az eljárás 0,93/0,78-os szenzitivitás/specificitás értéket ért el. Mivel a két adatféleség független, de egymást kiegészítő információt hordozott a vizsgálat várható kimenetére vonatkozóan, a három módszertanból ez bizonyult a legeredményesebbnek közelítve azt a minimum értéket, melyet a British Diabetic Association a szűrőprogramok számára meghatározott (legalább 0,80 szenzitivitás és 0,95 specificitás).

6.2. Angol nyelvű összefoglaló

Image processing and tear fluid proteomics based methods in the screening of diabetic retinopathy

Diabetic retinopathy (DR) is one of the most common complications of diabetes mellitus (DM), accounting for about 5 % of world blindness. Approximately 80 % of all patients with type 2 DM duration of at least 10 years suffer from some degree of DR. During the development of DR, patients may not notice changes of their vision and DR might be very advanced by the time patients have visual complaints and experience visual loss eventually. Timely diagnosis and therapy however can significantly decelerate its progress, necessitating regular DR screening or appropriate follow-up in all patients with DM. Involvement of human graders is currently universal and the less automation is in the screening system the greater the costs. Therefore our aim was to develop a novel automated method for DR screening.

In the first phase of the project our aim was to develop a novel method for DR screening based on the examination of tear fluid proteomics biomarker changes. We applied machine learning algorithms to predict whether the given patient with DM suffers from DR or not, based on the global proteomics pattern changes of their tear fluid. In the study we concluded that, due to its low sensitivity (0.74) and specificity (0.48) values, the proteomics based screening method alone is not appropriate for clinical application at its present format. However, in combination with other methods, it is able to improve the performance of the other methodology.

In the second phase of the study we used tear fluid proteomics based methodologies with retinal image processing combined in one single system. The image processing based method alone resulted in 0.84 sensitivity, and 0.81 specificity. Using the proteomics data for analysis 0.87 sensitivity, and 0.68 specificity values were achieved. The combined data analysis integrated the features of the proteomics data along with the number of detected MAs in the associated image, and achieved sensitivity/specificity values of 0.93/0.78. As the two different types of data represent independent and complementary information on the outcome, the combined model resulted in a reliable screening method that is comparable to the requirements of DR screening programs applied in clinical routine as specified by the British Diabetic Association (at least 0.80 sensitivity, and 0.95 specificity).

7. Irodalomjegyzék

7.1. Az értekezésben hivatkozott közlemények listája

1. Nemeth J, Frigýik A, Vastag O, Gocze P, Peto T: **Vaksági okok Magyarországon 1996 és 2000 között.** Szemészet. 2005, 142:127-133.
2. Shaw JE, Sicree RA, Zimmet PZ: **Global estimates of the prevalence of diabetes for 2010 and 2030.** Diabetes Res Clin Pract. 2010, 87(1):4-14.
3. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Lin JK, Farzadfar F, Khang YH, Stevens GA, Rao M, Ali MK, Riley LM, Robinson CA, Ezzati M: **National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants.** Lancet. 2011, 378(9785):31-40.
4. Morrish NJ, Wang SL, Stevens LK, Fuller JH, Keen H: **Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes.** Diabetologia. 2001, 44(2):14-21.
5. Nathan DM, Cleary PA, Backlund JY, Genuth SM, Lachin JM, Orchard TJ, Raskin P, Zinman B: **Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes.** The New England Journal of Medicine. 2005, 353(25):2643-53.

6. The Diabetes Control and Complications Trial Research Group: **The effect of intensive diabetes therapy on the development and progression of neuropathy.** Annals of Internal Medicine. 1995, 122(8):561-8.
7. Washington RE, Andrews RM, Mutter RL: **Emergency department visits for adults with diabetes, 2010.** HCUP Statistical Brief. 2013, 167.
8. Berta A, Hatvani I, Imre Gy, Kolozsvári L, Süveges I: **Szemészet.** Medicina Könyvkiadó Rt. 1998, 11:272-75.
9. Ashton N: **Studies of the retinal capillaries in relation to diabetic and other retinopathies.** Br J Ophthalmol. 1963, 47:521-38.
10. Bloodworth JMB: **Diabetic retinopathy.** Diabetes. 1962, 11:1-22.
11. Cogan DG, Toussaint D, Kuwabara T: **Retinal vascular patterns IV. Diabetic retinopathy.** Arch Ophthalmol. 1961, 66:366-78.
12. **Visual impairment and blindness Fact Sheet.** <http://www.who.int/mediacentre/factsheets/fs282/en/>
13. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, Chen SJ, Dekker JM, Fletcher A, Grauslund J, Haffner S, Hamman RF, Ikram MK, Kayama T, Klein BE, Klein R, Krishnaiah S, Mayurasakorn K, O'Hare JP, Orchard TJ, Porta M, Rema M, Roy MS, Sharma T, Shaw J, Taylor H, Tielsch JM, Varma R, Wang JJ, Wang N, West S, Xu L, Yasuda M, Zhang X, Mitchell P, Wong TY: **Global prevalence and major risk factors of diabetic retinopathy.** Diabetes Care. 2012; 35(3):556-564.
14. Kempen J H, O'Colmain BJ, Leske MC, Haffner SM, Klein R, Moss SE, Taylor HR, Hamman RF: **The prevalence of diabetic retinopathy among adults in the United States.** Arch Ophthalmol. 2004, 122(4):552-63.
15. Kovacs I, Salacz Gy: **A diabétesz szemészeti szövődményei.** Hippocrates. 2003, 5(5):290-292.
16. Wild S, Roglic G, Green A, Sicree R, King H: **Global prevalence of diabetes: Estimates for the year 2000 and projections for 2030.** Diabetes Care. 2004, 27(5):1047-53.
17. Larsen M, Godt J, Larsen N, Lund-Andersen H, Sjølie AK, Agardh E, Kalm H, Grunkin M, Owens DR: **Automated detection of fundus photographic red lesions in diabetic retinopathy.** Invest Ophthalmol Vis Sci. 2003, 44(2):761-6.
18. Cheung N, Mitchell P, Wong TY: **Diabetic retinopathy.** Lancet. 2010, 376(9735):124-136.
19. American Academy of Ophthalmology: **Preferred practice pattern guidelines: diabetic retinopathy.** 2008.
20. Facey K, Cummins E, Macpherson K, Morris A, Reay L, Slattery J: **Organisation of services for diabetic retinopathy screening.** Health Technology Assessment Report. 2002.
21. Olson JA, Strachan FM, Hipwell JH, Goatman KA, McHardy KC, Forrester JV, Sharp PF: **A comparative evaluation of digital imaging, retinal photography and optometrist examination in screening for diabetic retinopathy.** Diabet Med. 2003, 20(7):528-34.
22. Gibbins RL, Owens DR, Allen JC, Eastman L: **Practical application of the European Field Guide in screening for diabetic retinopathy by using ophthalmoscopy and 35 mm retinal slides.** Diabetologia. 1998, 41(1):59-64.

23. Sundling V, Gulbrandsen P, Straand J: **Sensitivity and specificity of Norwegian optometrists' evaluation of diabetic retinopathy in single-field retinal images, a cross-sectional experimental study.** BMC Health Serv Res. 2013, 13:17.
24. O'Hare JP, Hopper A, Madhavan C, Charny M, Purewell TS, Harney B, Griffiths J: **Adding retinal photography to screening for diabetic retinopathy: a prospective study in primary care.** BMJ. 1996, 312(7032):679-682.
25. **Screening for Diabetic Retinopath.**
<http://www.mrcophth.com/focus1/Screening%20for%20Diabetic%20Retinopathy.htm>
26. Harding S, Garvican L, Talbot J: **The impact of national diabetic retinopathy screening on ophthalmology: the need for urgent planning.** Eye (Lond). 2005, 19(9):1009-1011.
27. Silva PS, Cavallerano JD, Aiello LM, Aiello LP: **Telemedicine and diabetic retinopathy: moving beyond retinal screening.** Arch Ophthalmol. 2011, 129(2):236-242.
28. Bragge P, Gruen RL, Chau M, Forbes A, Taylor HR: **Screening for presence or absence of diabetic retinopathy: a meta-analysis.** Arch Ophthalmol. 2011, 129(4):435-444.
29. Gibson OR, Segal L, McDermott RA: **A simple diabetes vascular severity staging instrument and its application to a Torres Strait Islander and Aboriginal adult cohort of north Australia.** BMC Health Serv Res. 2012, 12:185.
30. Sotland GS, Philip S, Fleming AD, Goatman KA, Sharp PF, McNamee P, Prescott GJ, Fonseca S, Olson JA: **Manual vs. automated: the diabetic retinopathy screening debate.** Ophthalmol Times. 2008, 4:2.
31. Bouhaimed M, Gibbins R, Owens D: **Automated detection of diabetic retinopathy: results of a screening study.** Diabetes Technol Ther. 2008, 10(2):142-148.
32. Fleming AD, Goatman KA, Philip S, Prescott GJ, Sharp PF, Olson JA: **Automated grading for diabetic retinopathy: a large-scale audit using arbitration by clinical experts.** Br J Ophthalmol. 2010, 94(12):1606-1610.
33. Niemeijer M, van Ginneken B, Staal J, Suttorp-Schulten MSA, Abramoff MD: **Automatic detection of red lesions in digital color fundus photographs.** IEEE Transaction on medical imaging. 2005, 24(5).
34. Niemeijer M, van Ginneken B, Russell SR, Suttorp-Schulten MS, Abramoff MD: **Automated detection and differentiation of Drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis.** Invest Ophthalmol Vis Sci. 2007, 48(5):2260-7.
35. Bouhaimed M, Gibbins R, Owens D: **Automated detection of diabetic retinopathy: results of a screening study.** Diabetes Technol Ther. 2008, 10(2):142-148.
36. Fleming AD, Goatman KA, Philip S, Prescott GJ, Sharp PF, Olson JA: **Automated grading for diabetic retinopathy: a large-scale audit using arbitration by clinical experts.** Br J Ophthalmol. 2010, 94(12):1606-1610.

37. Larsen N, Godt J, Grunkin M, Lund-Andersen H, Larsen M: **Automated detection of diabetic retinopathy in a fundus photographic screening population.** Invest Ophthalmol Vis Sci. 2003, 44(2):767-771.
38. Niemeijer M, van Ginneken B, Cree MJ, Mizutani A, Quellec G, Sanchez CI, Zhang B, Hornero R, Lamard M, Muramatsu C, Wu X, Cazuguel G, You J, Mayo A, Li Q, Hatanaka Y, Cochener B, Roux C, Karray F, Garcia M, Fujita H, Abramoff MD: **Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs.** IEEE Trans Med Imaging. 2010, 29(1):185-195.
39. Green-Church KB, Nichols KK, Kleinholz NM, Zhang L, Nichols JJ: **Investigation of the human tear film proteome using multiple proteomic approaches.** Mol Vis. 2008, 14:456-470.
40. Grus FH, Augustin AJ: **High performance liquid chromatography analysis of tear protein patterns in diabetic and non-diabetic dry-eye patients.** Eur J Ophthalmol. 2001, 11(1):19-24.
41. Van Haeringen NJ: **Clinical biochemistry of tears.** Surv Ophthalmol. 1981, 26(2):84-96.
42. Csutak A, Silver DM, Tozser J, Steiber Z, Bagossi P, Hassan Z, Berta A: **Plasminogen activator inhibitor in human tears after laser refractive surgery.** J Cataract Refract Surg. 2008, 34(6):897-901.
43. de Souza GA, Godoy LM, Mann M: **Identification of 491 proteins in the tear fluid proteome reveals a large number of proteases and protease inhibitors.** Genome Biol. 2006, 7(8):R72.
44. Herber S, Grus FH, Sabuncuo P, Augustin AJ: **Changes in the tear protein patterns of diabetic patients using two-dimensional electrophoresis.** Adv Exp Med Biol. 2002, 506:623-626.
45. Molloy MP, Bolis S, Herbert BR, Ou K, Tyler MI, van Dyk DD, Willcox MD, Gooley AA, Williams KL, Morris CA, Walsh BJ: **Establishment of the human reflex tear two-dimensional polyacrylamide gel electrophoresis reference map: new proteins of potential diagnostic value.** Electrophoresis. 1997, 18(15):2811-2815.
46. Reitz C, Breipohl W, Augustin A, Bours J: **Analysis of tear proteins by one- and two-dimensional thin-layer isoelectric focusing, sodium dodecyl sulfate electrophoresis and lectin blotting. Detection of a new component: cystatin C.** Graefes Arch Clin Exp Ophthalmol. 1998, 236(12):894-899.
47. Fung KY, Morris C, Sathe S, Sack R, Duncan MW: **Characterization of the in vivo forms of lacrimal-specific proline-rich proteins in human tear fluid.** Proteomics. 2004, 4(12):3953-3959.
48. Grus FH, Podust VN, Bruns K, Lackner K, Fu S, Dalmaso EA, Wirthlin A, Pfeiffer N: **SELDI-TOF-MS ProteinChip array profiling of tears from patients with dry eye.** Invest Ophthalmol Vis Sci. 2005, 46(3):863-876.
49. Versura P, Nanni P, Bavelloni A, Blalock WL, Piazzini M, Roda A, Campos EC: **Tear proteomics in evaporative dry eye disease.** Eye (Lond). 2010, 24(8):1396-1402.
50. Zhou L, Huang LQ, Beuerman RW, Grigg ME, Li SF, Chew FT, Ang L, Stern ME, Tan D: **Proteomic analysis of human tears: defensin expression after ocular surface surgery.** J Proteome Res. 2004, 3(3):410-416.
51. Csosz E, Boross P, Csutak A, Berta A, Toth F, Poliska S, Torok Zs, Tozser J: **Quantitative analysis of proteins in the tear fluid of patients with diabetic retinopathy.** J Proteomics. 2012, 75(7):2196-2204.

52. Heise EA, Fort PE: **Impact of diabetes on alpha-crystallins and other heat shock proteins in the eye.** J Ocul Biol Dis Infor. 2011, 4(1-2):62-69.
53. Shitama T, Hayashi H, Noge S, Uchio E, Oshima K, Haniu H, Takemori N, Komori N, Matsumoto H: **Proteome profiling of vitreoretinal diseases by cluster analysis.** Proteomics Clin Appl. 2008, 2(9):1265-1280.
54. Klaassen I, Van Noorden CJ, Schlingemann RO: **Molecular basis of the inner blood-retinal barrier and its breakdown in diabetic macular edema and other pathological conditions.** Prog Retin Eye Res. 2013, 34:19-48.
55. Heise EA, Fort PE: **Impact of diabetes on alpha-crystallins and other heat shock proteins in the eye.** J Ocul Biol Dis Infor. 2011, 4(1-2):62-69.
56. Li ZZ, Lu XZ, Liu JB, Chen L: **Serum retinol-binding protein 4 levels in patients with diabetic retinopathy.** J Int Med Res. 2010, 38(1):95-9.
57. Tao Y, Lu Q, Jiang YR, Qian J, Wang JY, Gao L, Jonas JB: **Apelin in plasma and vitreous and in fibrovascular retinal membranes of patients with proliferative diabetic retinopathy.** Invest. Ophthalmol. Vis. Sci. 2010, 51(8): 4237-4242.
58. dell'Omo R, Semeraro F, Bamonte G, Cifariello F, Romano MR, Costagliola C: **Vitreous mediators in retinal hypoxic diseases.** Mediators of Inflammation. 2013, Article ID 935301.
59. Praidou A, Klangas I, Papakonstantinou E, Androudi S, Georgiadis N, Karakiulakis G, Dimitrakos S: **Vitreous and serum levels of platelet-derived growth factor and their correlation in patients with proliferative diabetic retinopathy.** Curr Eye Res. 2009, 34(2):152-61.
60. Hernández C, Ortega F, García-Ramírez M, Villarroel M, Casado J, García-Pascual L, Fernández-Real J M, Simó R: **Lipopolysaccharide-binding protein and soluble CD14 in the vitreous fluid of patients with proliferative diabetic retinopathy.** Retina. 2010, 30(2):345-52.
61. Takada M, Ban Y, Yamamoto G, Ueda T, Saito Y, Nishimura E, Fujisawa K, Koide R, Mizutani M, Kozawa T, Shiraishi Y, Bando Y, Tachikawa T, Hirano T: **Periostin, discovered by nano-flow liquid chromatography and mass spectrometry, is a novel marker of diabetic retinopathy.** Biochemical and biophysical research communications. 2010, 399:221-226.
62. Kim K, Kim SJ, Yu HG, Yu J, Park KS, Jang IJ, Kim Y: **Verification of biomarkers for diabetic retinopathy by multiple reaction monitoring.** J Proteome Res. 2010, 5;9(2):689-99.
63. Zhou L, Beuerman RW, Chan CM, Zhao SZ, Li XR, Yang H, Tong L, Liu S, Stern ME, Tan D: **Identification of tear fluid biomarkers in dry eye syndrome using iTRAQ quantitative proteomics.** J Proteome Res. 2009, 8(11):4889-905.
64. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlett-Jones M, He F, Jacobson A, Pappin DJ: **Multiplexed protein**

- quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents.** Mol Cell Proteomics. 2004, 3(12):1154-69.
65. DeSouza L, Diehl G, Rodrigues MJ, Guo J, Romaschin AD, Colgan TJ, Siu KW: **Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry.** J Proteome Res. 2005, 4(2):377-86.
 66. Torok Zs, Peto T, Csoz E, Tukacs E, Molnar AM, Maros-Szabo Zs, Berta A, Tozser J, Hajdu A, Nagy V, Domokos B, Csutak A: **Tear fluid proteomics multimarkers for diabetic retinopathy screening.** BMC Ophthalmology. 2013, 13(1):40.
 67. Berta A: **Standardization of tear protein determinations: the effects of sampling, flow rate, and vascular permeability.** In The precorneal tear film in health, disease, and contact lens wear. Dry Eye Institute. 1986, 418-435.
 68. Reitz C, Breipohl W, Augustin A, Bours J: **Analysis of tear proteins by one- and two-dimensional thin-layer isoelectric focusing, sodium dodecyl sulfate electrophoresis and lectin blotting. Detection of a new component: cystatin C.** Graefes Arch Clin Exp Ophthalmol. 1998, 36(12):894-899.
 69. Karatzoglou A, Hornik K, Smola A, Zeileis A: **Kernlab—An S4 package for kernel methods in R.** J Stat Softw. 2004, 11:9.
 70. **R: A language and environment for statistical computing.** <http://www.r-project.org/>
 71. Cree MJ, Olson JA, McHardy KC, Sharp PF, and Forrester JV. **A fully automated comparative microaneurysm digital detection system.** Eye. 1997, 11:622–628.
 72. Quéllec G, Lamard M, Josselin PM, Cazuguel G, Cochener B, and Roux C: **Optimal wavelet transform for the detection of microaneurysms in retina photographs.** IEEE Trans Med Imaging. 2008, 27(9):1230–1241.
 73. Vincent L: **Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms.** IEEE Trans Im Proc. 1993, 2:176–201.
 74. Friedman JH: **Greedy function approximation: A gradient boosting machine.** Annals of Statistics. 2001, 29(5):1189-1232.
 75. **Generalized boosted regression models.** <http://cran.r-project.org/web/packages/gbm/index.html>
 76. Li B, Sheng M, Xie L, Liu F, Yan G, Wang W, Lin A, Zhao F, Chen Y: **Tear proteomic analysis of patients with type 2 diabetes and dry eye syndrome by two-dimensional-nano-liquid chromatography coupled with tandem mass spectrometry.** Invest Ophthalmol Vis Sci. 2014, 55(1):177-186.
 77. **Smartbiobank.** <http://smartbiobank.com/>
 78. **R: A language and environment for statistical computing.** <http://www.r-project.org/>
 79. **GNU Octave.** <http://www.gnu.org/software/octave/>
 80. **23 and Me.** <https://www.23andme.com/howitworks/>
 81. White FM: **The potential cost of high-throughput proteomics.** Sci Signal. 2011, 4(160):pe8.

82. Abramoff MD, Reinhardt JM, Russell SR, Folk JC, Mahajan VB, Niemeijer M, Quéllec G: **Automated early detection of diabetic retinopathy**. Ophthalmology. 2010, 117(6):1147-54.
83. Goatman K, Charnley A, Webster L, Nussey S: **Assessment of automated disease detection in diabetic retinopathy screening using two-field photography**. PLoS One. 2011; 6(12):e27524.
84. Torok Zs, Peto T, Csosz E, Tukacs E, Molnar A, Berta A, Tozser J, Hajdu A, Nagy V, Domokos B, Csutak A: **Combined methods for diabetic retinopathy screening, using retina photographs and tear fluid proteomics biomarkers**. Journal of Diabetes Research, "Complications of Diabetes" (Special issue). 2014, Article ID 623619. <http://dx.doi.org/10.1155/2014/623619> In press.
85. Kuncheva L, Whitaker C: **Measures of diversity in classifier ensembles, and their relationship with the ensemble accuracy**. Machine Learning. 2003, 51(2):181-207.
86. Sollich P, Krogh A: **Learning with ensembles: How overfitting can be useful, Advances in Neural Information Processing Systems**. Advances in Neural Information Processing Systems. 1996, 8:190-196.
87. Lee JM, McMahon PM, Kong CY, Kopans DB, Ryan PD, Ozanne EM, Halpern EF, Gazelle GS: **Cost-effectiveness of breast MR imaging and screen-film mammography for screening BRCA1 gene mutation carriers**. Radiology. 2010, 254(3):793-800.

7.2. Az értekezés alapjául szolgáló saját közlemények listája

1. Csosz E, Boross P, Csutak A, Berta A, Toth F, Poliska S, **Torok Zs**, Tozser J: Quantitative analysis of proteins in the tear fluid of patients with diabetic retinopathy. J Proteomics. 2012, 75(7):2196-2204. Journal of Proteomics. 75 (7), 2196-2204, 2012.

DOI: <http://dx.doi.org/10.1016/j.jprot.2012.01.019>

IF:4.088

2. **Torok Zs**, Peto T, Csosz E, Tukacs E, Molnar AM, Maros-Szabo Zs, Berta A, Tozser J, Hajdu A, Nagy V, Domokos B, Csutak A: Tear fluid proteomics multimarkers for diabetic retinopathy screening.

BMC Ophthalmol. 13 (40), 8 p., 2013.

DOI: <http://dx.doi.org/10.1186/1471-2415-13-40>

IF:1.075

3. **Torok Zs**, Peto T, Csosz E, Tukacs E, Molnar AM, Berta A, Tozser J, Hajdu A, Nagy V, Domokos B, Csutak A: Combined methods for diabetic retinopathy screening, using retina photographs and tear fluid proteomics biomarkers.

Journal of Diabetes Research.

2014, Article ID 623619. In press.

IF:3.536

A disszertáció alapjául szolgáló közlemények összesített IF: 8.699

8. Tárgyszavak

8.1. Magyar nyelvű tárgyszavak

RD szűrés, könny proteomika, képfeldolgozás, digitális retinafotózás, gépi tanulás, biomarker, MA, könny fehérje koncentráció.

8.2. Angol nyelvű tárgyszavak

Diabetic retinopathy screening, tear fluid proteomics, image processing, digitalized retina photography, machine learning, biomarker, microaneurysm, tear fluid protein concentration.

9. Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőmnek, Dr. Csutak Adrienne-nek, akitől munkám során végig rengeteg szakmai útmutatást és gyakorlati segítséget kaptam. Szeretném kifejezni hálámat Dr. Csősz Évának a könnyminták feldolgozásáért, a proteomikai vizsgálatok elvégzéséért és az eredmények interpretálása során nyújtott segítségéért. Dr. Tőzsér József Professzor Úrnak és a Biokémiai és Molekuláris Biológiai Intézet, valamint a Proteomikai és Szolgáltató Laboratórium dolgozóinak, hogy segítették munkámat. Köszönöm Dr. Berta András Professzor Úrnak, hogy lehetővé tette számomra a Szemklinikán folyó kutatási tevékenységbe való részvételt és támogatott abban, hogy jelen munka célkitűzéseit megvalósíthassuk. A könny-mintavétel és a fotódokumentáció elkészítésében nyújtott segítséget köszönöm Szabó Katalinnak, Pető Tündének és a Szemklinikai többi dolgozójának. Külön köszönöm Tukacs Editnek a vizsgálatok során végzett biostatistikai elemzések és matematikai modellek alkalmazása, valamint az eredmények interpretálása során nyújtott szakmai támogatást és gyakorlati segítséget. Köszönöm Domokos Bálintnak, Balla Anettnek és az AstridBio Kft. többi munkatársának, hogy a képfeldolgozási és gépi tanulás alapú technológiák alkalmazásával segítettek munkámat. Dr. Pető Tündének köszönöm, hogy a retinopathia diabetica szűrés szakértőjeként segítette munkámat és betekintést engedett egy a nemzetközi élvonalba tartozó szűrőcentrum, a londoni Moorfields Eye Hospital szakmai tevékenységébe. Dr. Hajdú Andrásnak köszönöm, hogy szakértelmével és munkájával a képfeldolgozás témájához kapcsolódó számos közösen megvalósított kutatási projekt során hozzájárult jelen munka megalapozásához.

Nem utolsó sorban szeretnék köszönetet mondani azoknak, akik mentorként támogatták a klinikai orvostudományok területén folytatott kutatási tevékenységemet. Dr. Kozlovsky Bertalan Főorvos Úrnak köszönöm azt, hogy sebész, érsebész szakorvosként és szakmai vezetőként klinikai szemléletét adta át nekem és támogatta azt, hogy klinikusként bekapcsolódhassak a Debreceni Egyetemen folyó kutatási

tevékenységbe. Dr. Muszbek László Professor Úrnak köszönhetően ismerhettem meg a molekuláris biológiai vizsgáló módszereket, támogatta és irányította a kutatás területén tett kezdeti lépéseimet. Dr. Hunyadi János Professor Úrnak, akitől az érbetegségek okozta bőrgyógyászati kórképeket és a kapcsolódó vizsgálati módszertanokat ismerhettem meg;

Szeretném megköszönni családomnak, hogy munkám során mindvégig mellettem állt és biztosította számomra azt a családi háttérrel, mely nélkül ez a munka nem születhetett volna meg.

10. Függelék